

Aspekte der Validierung eines Tests zur Kompetenz in Biologie

*Eine Studie zur Kompetenz in Biologie und ihren Teildimensionen
Konzept- und Prozesswissen*

Dissertation zur Erlangung des akademischen Grads Dr. phil. im Fach Erziehungswissenschaft

Verteidigt am 19.12.2012

an der Philosophischen Fakultät IV - Institut für Erziehungswissenschaften der Humboldt Universität
zu Berlin

von Dipl.-Päd. Nele Nicole Kampa

Präsident der Humboldt Universität zu Berlin

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Fakultät

Prof. Dr. Dr. Jürgen van Buer

Gutachter/Gutachterinnen: 1. Prof. Dr. Olaf Köller

2. Prof. Dr. Hans Anand Pant

3. Prof. Dr. Anette Upmeier zu Belzen

Zusammenfassung

Dimensionalitäts- und Zusammenhangsanalysen mit disziplinübergreifenden Kompetenzen haben durch den zunehmenden Einsatz von nationalen und internationalen Kompetenztestungen und deren wachsender Bedeutung im 21. Jahrhundert an Wichtigkeit zugenommen. Insbesondere für die naturwissenschaftlichen Disziplinen Biologie, Chemie und Physik liegen kaum Untersuchungen dieser Art vor.

Bezugnehmend auf die Kompetenzbereiche Fachwissen und Erkenntnisgewinnung der Bildungsstandards in den Naturwissenschaften werden im Rahmen des Begriffs *scientific literacy* die Dimensionen Konzept- und Prozesswissen von Biologie erarbeitet und einer Prüfung der internen Dimensionalität unterzogen. Die hierfür verwendeten Daten stammen von 3 156 Schülerinnen und Schülern der Jahrgangsstufe 10 aus acht Bundesländern. Zur Beantwortung der Forschungsfragen zu den internen Zusammenhängen und den Zusammenhängen mit den externen disziplinübergreifenden Konstrukten und Maßen sprachliche Kompetenzen, kognitive Grundfähigkeit, Fachnoten in unterschiedlichen akademischen Disziplinen sowie dem akademischen Selbstkonzept in Biologie, Chemie und Physik wurden entsprechende Strukturmodelle berechnet. Es konnte gezeigt werden, dass Kompetenz in Biologie aus den zwei Dimensionen Konzept- und Prozesswissen zusammengesetzt ist; diese korrelieren jedoch hoch miteinander.

Diese zweidimensionale Kompetenz in Biologie wird mit den beiden disziplinübergreifenden Konstrukten kognitive Grundfähigkeit und sprachliche Kompetenzen in Beziehung gesetzt. Neben den erwarteten hohen Zusammenhängen mit den Dimensionen von Kompetenz in Biologie lagen differentielle Effekte nur für Konzeptwissen vor. Sprachliche Kompetenzen hängen höher mit Konzeptwissen zusammen als Prozesswissen. Die Zusammenhänge mit den externen Kriterien Fachnoten und akademisches Selbstkonzept ergeben ein uneinheitliches Bild. Die Fachnoten zeigten keine der vermuteten distinkten Zusammenhangsmuster. Das Selbstkonzept in Biologie hängt mit beiden Dimensionen von Kompetenz in Biologie gleichermaßen zusammen. Die Selbstkonzepte in Chemie und Physik leisten kaum einen Erklärungsbeitrag.

Implikationen, Perspektiven sowie Grenzen der Ergebnisse – insbesondere bezogen auf die vorgefundene Zweidimensionalität von Kompetenz in Biologie – werden vor dem Hintergrund von Lehr-Lernbedingungen und zukünftiger Forschungsfragen diskutiert.

Schlüsselwörter

Naturwissenschaften, *scientific literacy*, Biologie, Bildungsstandards, Kompetenzen, Schulnoten, Selbstkonzept

Abstract

With the growing popularity of national and international competence testing in the 21st century, within-dimensionality analyses of competencies and relational studies with overarching disciplines have become quite important issues. Especially in Germany, where competence testing is a hotly debated topic, there is a need for more research on the newly developed test instruments, in particular in the domains chemistry, physics and biology.

Referring to the competence fields of *knowledge of facts* and *scientific inquiry* of the German National Educational Standards in Germany, the dimensions *concept* and *process knowledge* in biology are defined within the framework of scientific literacy. These dimensions are tested for dimensionality. The analyses are based on data from 3 165 students in eight German states. To answer the research questions on internal dimensionality and the relations to the external overarching constructs and measures structural models incorporating linguistic competencies, cognitive ability, grades in various academic domains and academic self-concepts in biology, chemistry and physics were applied. The results show that competence in biology consists of two highly correlated dimensions: concept- and process knowledge.

The study furthermore analyses the relationship between the two-dimensional competence of biology and the overarching constructs of linguistic competencies and cognitive ability. Besides the expected high relations with both dimensions of competence in biology, differential effects could be detected for concept knowledge. Linguistic competencies relate stronger to concept knowledge than cognitive abilities. The relation to the external measures of grades and academic self-concept are inconsistent. Concerning the grades, none of the expected patterns were found. Self-concept in biology relates in the same manner for both dimensions. Self-concept in chemistry and physics do not explain further variance in concept and process knowledge.

The implications, perspectives and limitations of the results – especially focusing on the revealed two-dimensionality of competence in biology – are discussed with regard to teaching and learning as well as future research.

Keywords

Science, scientific literacy, Biology, National Educational Standards, Competencies, Grades, Self-concept

Inhaltsverzeichnis

Hinführung	1
1 Fachdidaktische und bildungspolitische Verortung von naturwissenschaftlicher Kompetenz und Kompetenz in Biologie.....	4
1.1 Definition des Begriffs fachübergreifende Kompetenz.....	7
1.2 Beschreibung und Definition von naturwissenschaftlicher Kompetenz	9
1.2.1 Theoretische Annäherungen an den Begriff <i>scientific literacy</i>	11
1.2.2 Konzeptualisierung und Operationalisierung naturwissenschaftlicher Kompetenz in <i>large-scale assessment</i> Studien sowie deren Ergebnisse.....	16
1.2.2.1 Konzeptualisierung naturwissenschaftlicher Kompetenz in <i>large-scale assessment</i> Studien	16
1.2.2.2 Operationalisierung naturwissenschaftlicher Kompetenz in internationalen <i>large-scale assessment</i> Studien.....	23
1.2.2.3 Ergebnisse zur Struktur naturwissenschaftlicher Kompetenz	26
1.3 Beschreibung und Definition von Kompetenz in Biologie.....	32
1.3.1 Kompetenz in Biologie – Fakten- und Handlungswissen in der schulischen Praxis.....	33
1.3.1.1 Bildungsstandards im Fach Biologie.....	34
1.3.1.2 Schulcurricula am Beispiel der Bundesländer Berlin und Bayern	37
1.3.1.3 Von <i>scientific literacy</i> zu einer Definition von Kompetenz in Biologie	42
2 Grundlagen zur Validierung eines Tests zu Kompetenz in Biologie	45
2.1 Die Naturwissenschaften aus kognitionspsychologischer Perspektive.....	49
2.1.1 Deklaratives Wissen	52
2.1.2 Prozedurales Wissen	53
2.1.3 Das Zusammenspiel von prozeduralem und deklarativem Wissen	56
2.1.4 Erkenntnisgewinnung in den Naturwissenschaften und in der Biologie	61
2.1.4.1 Erkenntnisgewinnung als biologiespezifisches prozedurales Wissen.....	63
2.1.4.2 Erkenntnisgewinnung als biologiespezifisches Problemlösen.....	65
2.1.4.3 Synthese von Erkenntnisgewinnung als biologiespezifisches prozedurales Wissen und als biologiespezifisches Problemlösen	69

2.2	Die Naturwissenschaften aus differentialpsychologischer Perspektive.....	71
2.2.1	Zusammenhänge naturwissenschaftlicher Kompetenz mit verwandten kognitiven Konstrukten.....	72
2.2.1.1	Der Zusammenhang mit kognitiver Grundfähigkeit	73
2.2.1.2	Der Zusammenhang mit sprachlichen Kompetenzen.....	79
2.2.2	Zusammenhänge naturwissenschaftlicher Kompetenz mit Fachnoten und Persönlichkeitsmerkmalen.....	81
2.2.2.1	Der Zusammenhang mit Fachnoten in naturwissenschaftlichen und nicht-naturwissenschaftlichen Disziplinen	82
2.2.2.2	Der Zusammenhang mit Persönlichkeitsmerkmalen – Die naturwissenschaftlichen Selbstkonzepte.....	85
3	Fragestellungen zu Aspekten der Konstruktvalidität von Kompetenz in Biologie	91
3.1	Fragestellung zum Zusammenhang des Konzept- und Prozesswissens in Biologie – Fachdidaktisch-kognitionspsychologische Perspektive.....	91
3.2	Fragestellungen zum Zusammenhang von Kompetenz in Biologie mit verwandten kognitiven Konstrukten – Differentialpsychologische Perspektive I.....	93
3.3	Fragestellungen zum Zusammenhang von Kompetenz in Biologie mit Fachnoten und naturwissenschaftlichen Selbstkonzepten – Differentialpsychologische Perspektive II	96
4	Methode.....	100
4.1	Grundgesamtheit und Stichprobenziehung	100
4.2	Durchführung der Studie	102
4.3	Operationalisierung der Konstrukte.....	105
4.3.1	Testaufgaben zum Konzept- und Prozesswissen in Biologie	105
4.3.1.1	Entwicklung der Testaufgaben zum Konzept- und Prozesswissen	106
4.3.1.2	Beschreibung der Aufgabenmerkmale	107
4.3.2	Instrumente zur Validierung der Kompetenzmessung	116
4.4	Statistische Analysen	120
4.4.1	Die Modellierung der Messmodelle	120
4.4.2	Die Modellierung der Strukturmodelle.....	122

5	Ergebnisse	125
5.1	Hypothesentestung.....	127
5.1.1	Forschungshypothese A: Zusammenhang des Konzept- und Prozesswissens in Biologie – Fachdidaktisch-kognitionspsychologische Perspektive	127
5.1.2	Forschungshypothesen B1 sowie B2a und B2b: Zusammenhang von Kompetenz in Biologie mit verwandten kognitiven Konstrukten – Differentialpsychologische Perspektive I.....	132
5.1.3	Forschungshypothesen C und D: Zusammenhang von Kompetenz in Biologie mit Fachnoten und naturwissenschaftlichen Selbstkonzepten – Differentialpsychologische Perspektive II	134
6	Diskussion.....	141
6.1	Forschungshypothese A: Zusammenhang des Konzept- und Prozesswissens in Biologie ..	142
6.2	Forschungshypothesen B1 sowie B2a und B2b: Zusammenhang von Kompetenz in Biologie mit verwandten kognitiven Konstrukten.....	146
6.3	Fragestellung C: Zusammenhang von Kompetenz in Biologie mit Fachnoten in den naturwissenschaftlichen und nicht-naturwissenschaftlichen Fächern	150
6.4	Fragestellung D – Zusammenhang von Kompetenz in Biologie mit naturwissenschaftlichen Selbstkonzepten	153
6.5	Aspekte der Validierung eines Tests zur Kompetenz in Biologie – Übergreifende Diskussion, Grenzen und Ausblick.....	155
7	Anhang	159
8	Literaturverzeichnis	170

Tabellenverzeichnis

Tabelle 1: Inhaltliche Beschreibung der Dimensionen Prozesse und Wissen der naturwissenschaftlichen Kompetenz in der PISA-Studie 2006.....	18
Tabelle 2: Konzeptualisierung naturwissenschaftlicher Kompetenz der PISA-Studie 2006, der TIMS-Studie 1995/96, bei Duit et al. (2001) sowie bei Bybee (1997)	21
Tabelle 3: Verteilung der Aufgaben der internationalen Naturwissenschaftstests PISA 2006 und TIMSS 2007 auf Teilbereiche naturwissenschaftlicher Kompetenz	24
Tabelle 4: Messfehlerbereinigte Korrelationen zwischen den naturwissenschaftlichen Teilkompetenzen der PISA-Studie 2006 (OECD, 2009)	30
Tabelle 5: Inhalte der vier Kompetenzbereiche der Bildungsstandards im Fach Biologie	35
Tabelle 6: Exemplarische inhaltliche Ausgestaltung der Standards des Kompetenzbereichs Fachwissen nach Basiskonzepten	36
Tabelle 7: Exemplarische inhaltliche Ausgestaltung der Standards des Kompetenzbereichs Erkenntnisgewinnung	37
Tabelle 8: Mittlere Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für das Basiskonzept Entwicklung	38
Tabelle 9: Mittlere Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für den Kompetenzbereich Erkenntnisgewinnung	39
Tabelle 10: Kompetenzbezüge der Inhalte von Biologieunterricht im Berliner Schulcurriculum am Beispiel des Themenfeldes <i>Einheimische Pflanzen und ihre Bedeutung im Stoffkreislauf</i>	39
Tabelle 11: Unterrichtsinhalte im Fach Biologie der Jahrgangsstufe 10 im Bayerischen Curriculum...	41
Tabelle 12: Integration der Validitätsaspekte von Messick (1995) und der Quellen von Validität der AERA, APA & NCME (2004) sowie Einordnung der vorliegenden Arbeit.....	47
Tabelle 13: Messfehlerbereinigte Korrelationen der Naturwissenschaftstests mit den Lesetests der PISA Untersuchungen 2000, 2003, 2006 und 2009	80
Tabelle 14: Messfehlerbereinigte Korrelationen zwischen den Selbstkonzepten in Biologie, Chemie und Physik.....	89
Tabelle 15: Vergleich der Biologie-, der Gesamtstichprobe und der eliminierten Fälle (Nicht-Biologie) anhand des Alters und des Geschlechts (Geschlecht in Prozent).....	102
Tabelle 16: Zeitlicher Ablauf der Testung.....	103
Tabelle 17: Auszug aus dem Testdesign der Studie – Biologietesthefte.....	104
Tabelle 18: Skalenbeschreibung der Lückentests – entfernte Aufgaben, Schwierigkeitsverteilung der Testaufgaben sowie Modellfitwerte der beiden Messmodelle.....	117

Tabelle 19: Anzahl der Schülerinnen und Schüler mit Angaben zu den Schulnoten und zum Unterricht im Klassenverband pro akademische Disziplin (absolut, Klammern in Prozent).....	119
Tabelle 20: BIC der 1PL- und 2PL-Modelle für die der Untersuchung zu Grunde liegenden Konstrukte Konzept- und Prozesswissen, sprachliche Kompetenzen, kognitive Grundfähigkeit und akademisches Selbstkonzept	121
Tabelle 21: Gegenüberstellung der BIC der ein- und dreidimensionalen Analysen für die naturwissenschaftlichen Selbstkonzepte	124
Tabelle 22: Durchschnittliche Fachnote der Schülerinnen und Schüler in den naturwissenschaftlichen Disziplinen sowie in Deutsch und Mathematik	125
Tabelle 23: Produkt-Moment-Interkorrelationen der Summenwerte des Lückentests, KFT, Selbstkonzepts und der Fachnoten (Validierungsinstrumente)	126
Tabelle 24: Modellfitwerte und -tests zur vergleichenden Bewertung der drei 1PL-Modelle zu Kompetenz in Biologie	130
Tabelle 25: Messfehlerbereinigte Korrelationen zwischen den Dimensionen Konzept- und Prozesswissen mit den disziplinübergreifenden Konstrukten sprachliche Kompetenzen und kognitive Grundfähigkeit	132
Tabelle 26: Messfehlerbereinigte Korrelationen zwischen den Dimensionen Konzept- und Prozesswissen und den Fachnoten in Deutsch und Mathematik sowie den drei naturwissenschaftlichen Fächern Biologie, Chemie und Physik.....	135
Tabelle 27: Regressionsgewichte der Fachnoten Biologie, Chemie, Physik, Mathematik und Deutsch für die beiden Teildimensionen Konzept- und Prozesswissen in Biologie und deren Standardfehler	137
Tabelle 28: Messfehlerbereinigte Korrelationen zwischen den Dimensionen Konzept- und Prozesswissen mit den Selbstkonzepten in Biologie, Chemie und Physik.....	138
Tabelle 29: Inhaltliche Beschreibung der Dimensionen naturwissenschaftlicher Kompetenz konzeptualisiert bei PISA 2006 (OECD, 2007; PISA-Konsortium Deutschland, 2007) ..	159
Tabelle 30: Beispielhafte inhaltliche Ausgestaltung der Basiskonzepte des Kompetenzbereichs Fachwissen im Fach Biologie (KMK, 2005a).....	160
Tabelle 31: Inhaltliche Ausformulierung des Kompetenzbereichs Erkenntnisgewinnung im Fach Biologie (KMK, 2005a)	160
Tabelle 32: Standards für den Kompetenzbereich Fachwissen (KMK, 2005a).....	161
Tabelle 33: Standards für den Kompetenzbereich Erkenntnisgewinnung (KMK, 2005a)	162
Tabelle 34: Mittlere Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für den Kompetenzbereich Fachwissen (SBJS, 2006).....	163

Tabelle 35: Einfache, mittlere und erweiterte Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für den Kompetenzbereich Fachwissen – Entwicklungs-Konzept	164
Tabelle 36: Messfehlerbereinigte Korrelationen der Fachnoten Biologie, Chemie, Physik, Deutsch und Mathematik	164
Tabelle 37: Regressionsgewichte der Fachnoten Chemie, Physik, Mathematik und Deutsch für die beiden Teildimensionen Konzept- und Prozesswissen in Biologie und deren Standardfehler	165
Tabelle 38: Regressionsgewichte der Teildimensionen Konzept- und Prozesswissen auf die Selbstkonzepte Chemie und Physik	165
Tabelle 39: Regressionsgewichte der Teildimensionen Konzept- und Prozesswissen in Biologie auf die Fachnoten Biologie, Chemie, Physik, Mathematik und Deutsch, Teilpopulationen Gymnasien ($n = 1\,496$) und Schulformen mit nicht-gymnasialen Bildungsgängen ($n = 1\,669$)	166

Abbildungsverzeichnis

Abbildung 1: Entwurf der Wissens-/Reasoningtypen und ihrer Beziehungen zueinander.....	14
Abbildung 2: Kompetenzstrukturmodell naturwissenschaftlicher Kompetenz	15
Abbildung 3: Das ACT-Modell von Anderson	51
Abbildung 4: Darstellung der Wissensstrukturen von Markowitsch.....	56
Abbildung 5: Iteratives Modell des Erlernens von Konzept- und Prozesswissen.....	60
Abbildung 6: Rahmenkonzept wissenschaftsmethodischer Kompetenzen von Mayer (2007)	67
Abbildung 7: Ein Modell für die Naturwissenschaften von Gott & Duggan (1995)	70
Abbildung 8: Die Drei-Stratum-Theorie von Carroll.....	75
Abbildung 9: Struktur des Selbstkonzepts nach Marsh, Byrne und Shavelson.....	86
Abbildung 10: Graphische Darstellung der Forschungshypothesen B1, B2a und B2b.....	95
Abbildung 11: Graphische Darstellung der Forschungshypothese C	97
Abbildung 12: Graphische Darstellung der Forschungshypothese D.....	98
Abbildung 13: Darstellung des Aufgabenentwicklungsprozesses	107
Abbildung 14: Beispielaufgabe zum Konzeptwissen in Biologie	109
Abbildung 15: Anzahl vorgelegter Testaufgaben für den Kompetenzbereich Fachwissen.....	110
Abbildung 16: Schwierigkeitsverteilung der Testaufgaben zu Konzeptwissen in der Jahrgangsstufe 10	111
Abbildung 17: Beispielaufgabe zum Prozesswissen in der Biologie.....	114
Abbildung 18: Anzahl vorgelegter Testaufgaben für den Kompetenzbereich Erkenntnisgewinnung 115	
Abbildung 19: Schwierigkeitsverteilung der Testaufgaben zu Prozesswissen in der Jahrgangsstufe 10	115
Abbildung 20: Spezifikation der Strukturmodelle Globale Kompetenz in Biologie (Modell Aa), Konzept- und Prozesswissen in Biologie (Ab) und zwei Zufallsdimensionen von Kompetenz in Biologie (Ac).....	128
Abbildung 21: Zusammenhang der Dimensionen Konzept- und Prozesswissen von Kompetenz in Biologie	131
Abbildung 22: Ergebnisse des Strukturmodells zu den Forschungshypothesen B1, B2a und B2b – Konzept- und Prozesswissen und sprachliche Kompetenzen sowie kognitive Grundfähigkeit	133
Abbildung 23: Spezifikation des Strukturmodells zu den Forschungshypothesen C1, C2 und C3 – Kompetenzen in Biologie und Fachnoten in Biologie, Chemie und Physik sowie in Mathematik und Deutsch	136

Abbildung 24: Spezifikation der Strukturmodelle zur Forschungshypothese D – Naturwissenschaftliche Selbstkonzepte und Konzept- und Prozesswissen in Biologie	139
Abbildung 25: Ergebnisse der Strukturmodelle zu Forschungshypothese D – Naturwissenschaftliche Selbstkonzepte und Konzept- und Prozesswissen in Biologie	140
Abbildung 26: Messmodell für Biologie Konzeptwissen	166
Abbildung 27: Messmodell für Biologie Prozesswissen	166
Abbildung 28: Messmodell für kognitive Grundfähigkeit	167
Abbildung 29: Messmodell für sprachliche Kompetenzen – Lückentest 1	167
Abbildung 30: Messmodell für sprachliche Kompetenzen – Lückentest 2	168
Abbildung 31: Messmodell für Selbstkonzept in Biologie	168
Abbildung 32: Messmodell für Selbstkonzept in Chemie	168
Abbildung 33: Messmodell für Selbstkonzept in Physik	169

„Nur die Anwendung, nicht der Besitz, macht den Werth des Reichthums“ (Treviranus, 1802, S. III)

Danksagung

Diese Doktorarbeit wäre ohne die Unterstützung von vielen lieben Menschen nicht in der vorliegenden Qualität zu Stande gekommen. Daher möchte ich an dieser Stelle einigen Personen danken, die nicht unwesentlich am Gelingen dieser Arbeit beteiligt waren.

Prof. Köller, Prof. Upmeyer zu Belzen und Prof Pant danke ich für die Begutachtung der vorliegenden Arbeit. Meinen Betreuern Prof. Köller und Dr. Schroeders danke ich für ihre Förderung und Unterstützung während meiner Zeit als Doktorandin am Institut zur Qualitätsentwicklung im Bildungswesen (IQB). Ihre wertvollen Ratschläge bei den ungezählten Treffen und die wiederum ungezählten Worte, die sie in Vorversionen dieser Doktorarbeit kommentiert haben, haben mich wissenschaftlich an die Stelle gebracht, an der ich jetzt stehe. Neben dieser Unterstützung hat mich meine Zweitgutachterin, Prof. Upmeyer zu Belzen, mit Ihren Ratschlägen auf noch unbekannte Literaturpfade bringen können.

Neben der offiziellen Betreuung hat auch die inoffizielle Unterstützung innerhalb meines Teams und im Institut sowohl zu meiner wissenschaftlichen Weiterentwicklung als auch schlicht und ergreifend zu einer schönen Zeit geführt. Insbesondere danke ich Nicola Klebba und Michael Katzenbach für die unermüdliche Unterstützung und das Interesse an dieser Arbeit. Auch dem weiteren Nawi-Team danke ich für die Unterstützung und die vielen schönen persönlichen Momente von Kuchenrunden bis zu Gute-Nachtschichten. Es hat unwahrscheinlich viel Spaß gemacht mit Euch an diesem großen Projekt zu arbeiten. Des Weiteren danke ich den Teilnehmerinnen und Teilnehmern des internen Kolloquiums für die wertvollen Rückmeldungen über den gesamten Zeitraum und meinen Bürokolleginnen und –kollegen für die gemeinsam verbrachte, wundervolle Zeit.

Neben der Unterstützung von professioneller Seite habe ich auch viel Unterstützung von meiner Familie und meinen Freunden erhalten. Dafür möchte ich Euch an dieser Stelle danken. Im Besonderen danke ich meiner Mutter, Nicole Kampa, die die Fertigstellung der Doktorarbeit nicht mehr miterleben konnte. Sie wird in meinem Herzen weiterleben und daher widme ich ihr diese Arbeit. Für die Unterstützung in der schweren Zeit nach ihrem Tod danke ich jedem in meiner Familie. Ganz besonders danke ich meinem Vater, Andreas Kampa, für die wichtigen Entscheidungen, die er während meiner Doktorarbeitszeit getroffen hat. Nicht zuletzt danke ich meiner Tante, Yvette Herde, für die Unterstützung und die gemeinsamen Jahre.

Danke für all die persönliche und akademische Hilfe, die ich von Ihnen allen und Euch allen in den letzten Jahren erhalten habe. Es war eine wunderschöne und intensive Zeit.

Hinführung

Die Begriffe Kompetenzen und Kompetenzentwicklung haben sowohl in der Bildungsforschung als auch in der Bildungspolitik stark an Bedeutung gewonnen. Disziplinspezifische Kompetenzen – insbesondere in den Sprachen, der Mathematik und den Naturwissenschaften – sind in das Zentrum des bildungspolitischen Interesses gerückt. Mit diesem Bedeutungszuwachs ging ein Paradigmenwechsel von der Input- zur Outputorientierung in der deutschen Bildungslandschaft einher (Helmke, 2000). Als eine spezifische Konsequenz aus den Ergebnissen internationaler *large-scale assessment* Studien (Baumert, Bos & Lehmann, 2000; Deutsches PISA-Konsortium, 2001) für die Naturwissenschaften kann die Publikation der Bildungsstandards für die Fächer Chemie, Biologie und Physik angesehen werden (Kultusministerkonferenz [KMK], 2005a; KMK, 2005b; KMK, 2005c). Diese Standards sind an das international akzeptierte Konzept der *scientific literacy* angelehnt, welches aufgrund der geforderten Umfänglichkeit und des hohen Niveaus naturwissenschaftlicher Kompetenz durchaus kritisch betrachtet wird (Shamos, 1995). *Scientific literacy* beschreibt die naturwissenschaftlichen Kompetenzen, welche eine Person benötigt, um sowohl disziplinspezifische als auch naturwissenschaftsbezogene alltägliche Probleme bewältigen zu können (*Organisation for Economic Co-operation and Development* [OECD], 2000). Die naturwissenschaftliche Kompetenz befähigt Schülerinnen und Schülern dazu, erworbenes Wissen lebenslang flexibel im Alltag anzuwenden (OECD, 2007; PISA-Konsortium Deutschland, 2007). So wird durch den Kompetenzansatz – neben der reinen Fachwissenskomponente – zu erlernendes Wissen um Prozessaspekte wie Erkenntnisgewinnung, Bewertung oder Kommunikation (KMK, 2005a; KMK, 2005b; KMK, 2005c) sowie um soziale und psychologische Bereiche wie motivationale Einstellungen erweitert (Klieme & Leutner, 2006; Lau & Roeser, 2002; Weinert, 2001). Auch die verschiedenen Konzeptualisierungen von *scientific literacy* greifen diese Mehrdimensionalität und Erweiterungen auf. Sie beinhalten mehrere Teilbereiche wie konzeptuelles Wissen, Prozesswissen bzw. Urteils- und Entscheidungsfähigkeit (Bybee, 1997; Niedersächsisches Kultusministerium, 1995; OECD, 2000; OECD 2007) oder unterstreichen die Wichtigkeit der Verknüpfung der Motivationslagen mit der Entwicklung von *scientific literacy* (Shavelson et al., 2008).

Bei dem Kompetenzbegriff handelt es sich um ein mehrdimensionales Konstrukt, das unterschiedlichste Aspekte vereint. Bezogen auf eine Verortung naturwissenschaftlicher Kompetenz bleibt die Frage nach deren strukturellen Teilbereichen und Zusammenhängen. Kompetenzmessungen müssen die Binnenstruktur einer Kompetenz abbilden (Klieme & Leutner, 2006). Eine Betrachtung naturwissenschaftlicher Kompetenz aus unterschiedlichen Perspektiven legt

eine Unterteilung in zwei Facetten nahe. Demnach bestünde Kompetenz in den Naturwissenschaften aus einer Konzept- und einer Prozesswissenskomponente. Diese Unterteilung wird sowohl in theoretischen Konstrukten als auch in der Kompetenzmessung naturwissenschaftlicher Kompetenz bereits angewandt (z. B. Mullis et al., 2003; OECD, 2007; Shavelson, Young, Ayala, Brandon, Furtak & Ruiz-Promi, 2008). Der Prozessaspekt verhilft zu einem realistischen Blick auf die Eigenschaften der Naturwissenschaften. Die Naturwissenschaften nicht als harte Wissenschaft mit der Idee der absoluten naturwissenschaftlichen Objektivität zu verstehen zeigt den Schülerinnen und Schülern, dass sie gewonnene Daten und Informationen interpretieren müssen – auch als zukünftige Bürgerinnen und Bürger.

Validitätsprüfungen bezogen auf Kompetenzbereiche werden gefordert (Bernholt, Parchmann & Commons, 2009) und wurden bisher einzig anhand des nationalen Naturwissenschaftstests der *Programme for International Student Assessment* (PISA)-Studie 2003 vorgenommen (PISA-Konsortium Deutschland, 2004). Zudem sind die vorliegenden Kompetenzmodelle noch nicht hinreichend elaboriert. Ihre Bezüge zu didaktischen, erziehungswissenschaftlichen und psychologischen Modellen sowie zu Kompetenzen, wissenschaftlichen Konstrukten und im deutschen Kontext zu den Bildungsstandards sind noch nicht befriedigend verortet (Köller, 2008). Im vorliegenden Dissertationsprojekt werden diese Validitätsprüfung und die damit einhergehende Verortung anhand von Aufgaben zu den Bildungsstandards in den Naturwissenschaften angestrebt. Sie wird konkret für die akademische Disziplin Biologie durchgeführt. Der aktuelle Forschungsstand zu diesem Thema zeigt (s. Seite 33), dass mit diesem Vorhaben weitestgehend Neuland betreten wird.

Für die „Evaluation der Standards in den Naturwissenschaften für die Sekundarstufe I“ (ESNaS) werden am Institut zur Qualitätsentwicklung im Bildungswesen (IQB) Testaufgaben für Schülerinnen und Schüler der Jahrgangsstufe 9 und 10 entwickelt. Auf Grundlage einer im Herbst 2009 durchgeführten Studie mit 3 156 Schülerinnen und Schülern soll die Struktur von Kompetenz in Biologie bezogen auf die Teilbereiche Konzept- und Prozesswissen getestet werden. Die theoretischen Ausdifferenzierungen von *scientific literacy* und naturwissenschaftlicher Kompetenz lassen erwarten, dass eine empirische Prüfung das Bestehen der Dimensionen Konzept- und Prozesswissen bestätigt. Durch das Aufdecken der strukturellen Beschaffenheit und der Zusammenhänge von Teilbereichen wird es möglich sein, eine angepasste Kompetenzförderung und ein diesen Strukturen angemessenes Unterrichten voranzutreiben (Bayrhuber, 2007; Hamann, 2006). Des Weiteren werden Zusammenhänge mit den disziplinübergreifenden Konstrukten sprachliche Kompetenzen und kognitive Grundfähigkeit sowie mit unterschiedlichen Fachnoten und naturwissenschaftlichen Selbstkonzepten geprüft. Da der Kompetenzbegriff ein stark

anwendungsbezogenes Konstrukt mit vielfältigen Aspekten ist, wird davon ausgegangen, dass die disziplinübergreifenden Fähigkeiten einen bedeutenden Teil von Kompetenz in Biologie gestalten.

Im theoretischen Teil der Arbeit werden die Naturwissenschaften bzw. die Biologie aus drei verschiedenen Perspektiven betrachtet: der fachdidaktischen, der kognitionspsychologischen und der differentialpsychologischen Perspektive. Zu Beginn wird auf die fachdidaktische Perspektive eingegangen. In ihr werden die Bildungsstandards für das Fach Biologie, curriculare Aspekte und die Konzeptualisierungen der Naturwissenschaften in internationalen und nationalen *large-scale* assessment Studien aufgegriffen.

1 Fachdidaktische und bildungspolitische Verortung von naturwissenschaftlicher Kompetenz und Kompetenz in Biologie

Die Notwendigkeit von Kompetenzen in den Naturwissenschaften ist gegenüber der Notwendigkeit von Kompetenzen in den akademischen Disziplinen Mathematik und Deutsch weniger evident. Während geringe Mathematik- und Lesekompetenzen ganz unmittelbare Konsequenzen für die Teilhabe in modernen Gesellschaften haben, ist der diesbezügliche Wert der Naturwissenschaften nicht direkt ersichtlich (Baumert, Lehmann & Lehrke, 1997; Rost, Prenzel, Carstensen, Senkbeil & Groß, 2004; Shamos, 2002). Personen mit geringen mathematischen Kompetenzen oder mit geringer Lesekompetenz werden z. B. Probleme mit der Verwaltung ihres eigenen Vermögens haben. Fehlende Kenntnisse in den Naturwissenschaften führen zunächst nicht zu diesen offensichtlichen individuellen Defiziten in der elementaren Teilhabe am gesellschaftlichen Leben (Rost et al., 2004). Dennoch gilt es naturwissenschaftliche Entscheidungen zu treffen, z. B. ob und gegen welche Krankheiten das eigene Kind geimpft werden soll oder ob man sich einer bestimmten medizinischen Maßnahme unterzieht. Auf der gesellschaftlichen Ebene sind die Auswirkungen der Naturwissenschaften (z. B. Präventionsmaßnahmen einer Gemeinschaft gegen Naturgefahren) demgegenüber sichtbarer. Kompetenzen in den Naturwissenschaften stellen vielmehr ein Instrument zur kulturellen, wirtschaftlichen, gesellschaftlichen und politischen Teilhabe dar (Baumert et al., 1997; Gräber & Nentwig, 2002; Rost et al., 2004) und werden nicht zu den Schlüsselkompetenzen gerechnet (Weinert, 2001). Im Informations- und Technologiezeitalter mit seinen rapiden sozialen und technologischen Veränderungen (Roberts & Gott, 1999; Rychen, 2001) – insbesondere in den Biowissenschaften (Wong & Hodson, 2009) – steigt jedoch die Bedeutung von Fähigkeiten und Wissen im naturwissenschaftlichen Bereich. So nehmen das durch die Naturwissenschaften generierte Wissen und Kompetenzen in den Naturwissenschaften einen gehobenen Stellenwert in postindustriellen Gesellschaften ein (Campbell & Reece, 2003; Duit, Häußler & Prenzel, 2001; Oelkers, 2002; Shen, 1975), welcher in Zukunft noch größer werden wird (Shen, 1975). Die Naturwissenschaften werden gar zu den Haupterrungenschaften unserer Kultur gezählt (Duit et al., 2001; Gräber & Nentwig, 2002; Harlen, 2001; Osborne, 2007;).

In der heutigen Gesellschaft haben naturwissenschaftliche Thematiken direkten Einfluss auf das Leben der Bürgerinnen und Bürger. So sind viele Katastrophen in der modernen Welt (z. B.: Atomunglück in Fukushima) erst durch die Fortschritte in den Naturwissenschaften möglich geworden. Aber auch uns umgebende Phänomene wie der Umgang mit Krankheiten, Tod und Leben werden durch diese Fortschritte verändert und bedeutsamer. Dieser gehobene Stellenwert spiegelt

sich auch in der Rangliste der am stärksten besetzten Ausbildungsberufe der Fachserie 11 des Statistischen Bundesamtes wider. So stehen auf den ersten vier Rangplätzen der Männer die technisch geprägten Berufe Kraftfahrzeugmechatroniker, Industriemechaniker, Elektroniker und Anlagenmechaniker für Sanitär-, Heizungs- und Klimatechnik. Unter den ersten zwanzig Berufen können acht Berufe mit technischem oder naturwissenschaftlichem Hintergrund lokalisiert werden. In der Rangliste der Frauen befindet sich auf dem dritten Platz der Beruf medizinische Fachangestellte und auf dem siebten Platz die zahnmedizinische Angestellte (Destatis, 2011). Baumert (2002) nennt daher als einen von vier Modi der Weltbegegnungen, die unser „kanonisches Orientierungswissen“ (Baumert, 2002, S. 113) ausmachen, die kognitiv-instrumentelle Modellierung der Welt. In diesem Modus repräsentieren die Naturwissenschaften eine „dominante Form der Erschließung und Rekonstruktion der Welt“ (Baumert et al., 1997, S. 61). Um über die kognitiv-instrumentelle Modellierung die Welt begreifen zu können, benötigt eine Schülerin bzw. ein Schüler eine systematische Vermittlung von Kompetenzen in Mathematik und den Naturwissenschaften im schulischen Rahmen (Baumert et al., 2000).

Naturwissenschaftliche Kompetenzen müssen über das Wissen um naturwissenschaftliche Fakten und Konzepte hinausgehen. Sie umfassen auch Handlungsaspekte sowie soziale, motivationale und affektive Aspekte, die eine Person zur Teilhabe an der Gesellschaft sowohl auf der gesellschaftlichen als auch auf der individuellen Ebene befähigen. Auf der gesellschaftlichen Ebene wirkt naturwissenschaftliche Kompetenz auf die ökonomische Lage eines Landes und stellt einen starken Wirtschaftsfaktor dar (Gräber & Nentwig, 2002; Minkee, Lavonen & Ogawa, 2009; Winick, Avallone, Smith & Crovo, 2008). Sie führt des Weiteren zu einer kritisch reflektierten, unterstützenden Haltung und realistischen Erwartungen gegenüber den Naturwissenschaften (Bayrhuber, 2007; KMK, 2005a), einer gehobenen Position eines Staates in einer technologisierten Welt, einer kritisch positiv eingestellten Bevölkerungsmehrheit sowie zu einem bewussten, nachhaltigen Umgang mit der unmittelbaren Umgebung und den natürlichen Ressourcen eines Landes (KMK, 2005a).

Auf der individuellen Ebene begegnen Personen in vielfältigen Situationen naturwissenschaftlichen Thematiken wie z. B. der Gentechnik oder dem Klimawandel. Um sich mit diesen Themen auseinanderzusetzen, sind solide disziplinspezifische Kenntnisse von Nöten. Ohne diese Kenntnisse bleiben Partizipationsmöglichkeiten bezogen auf gesellschaftlich relevante Themen wie z. B. nachhaltige Entwicklung verborgen. Naturwissenschaftlich kompetente Personen können sich in einer von den Naturwissenschaften und Technik geprägten Welt sicher bewegen und verstehen ihr Funktionieren. Die gesellschaftlich relevanten Themen scheinen in den Unterricht zwar Eingang gefunden zu haben, allerdings ist (auch naturwissenschaftlichen) Lehrkräften der direkte Anwendungsbezug in der Lebensführung nicht immer bewusst (Rieß & Mischo, 2007).

Für diese beiden Wirkungsebenen sollten Personen einer Gesellschaft über das notwendige Werkzeug zur Bewältigung vielfältiger Situationen in Form von naturwissenschaftlicher Kompetenz verfügen. Die Naturwissenschaften durchziehen im aktuellen Jahrhundert alle Bereiche der Gesellschaft und bilden einen nicht vernachlässigbaren Teil unserer kulturellen Identität (Gräber & Nentwig, 2002; KMK, 2005a). Für ein Verständnis und die kritische Reflektion von Entwicklungen z. B. in den Biowissenschaften sollten Bürgerinnen und Bürger zunehmend technisierter Gesellschaften einen gemeinsamen Erfahrungs- und Wissenshintergrund im naturwissenschaftlichen Bereich und entsprechende Fähigkeiten und Fertigkeiten aufweisen (Hurd, 1998). So ist z. B. das Bilden einer wissenschaftlich fundierten Meinung im Zusammenhang mit politischen Debatten über den Klimawandel von Vorteil. Aus der Notwendigkeit naturwissenschaftlicher Fähigkeiten für eine aktive Teilnahme an gesellschaftlicher Kommunikation und Meinungsbildung ergibt sich folgende Frage: Wie kann naturwissenschaftliche Bildung Bürgerinnen und Bürger auf die zukünftige Teilhabe in Gesellschaften vorbereiten, wenn deren Wissensbasis einem ständigen Wandel unterzogen ist? An diese Frage schließt sich die Fragestellung an, welche Kompetenzen deutschen Schülerinnen und Schülern vermittelt werden müssen (Shavelson et al., 2008; Westby & Torres-Velásquez, 2000).

Die Ausführungen machen deutlich, dass naturwissenschaftliches Wissen kontinuierlich neu angeeignet werden muss. Daher sollte ein Ziel der Vermittlung naturwissenschaftlicher Kompetenz sein, alle Schülerinnen und Schüler mit einem Verständnis über die Naturwissenschaften und deren Methodik auszustatten (*American Association for the Advancement of Science* [AAAS], 1990; AAAS, 1993; Hodson, 2003; Hurd, 1998; Shen, 1975). Das im schulischen Kontext im Vordergrund stehende Vermitteln von Faktenwissen (Bybee, 2002; Hammann, 2006; Kattmann, 2003a) – also von den intellektuellen Produkten der Naturwissenschaften – scheint diesem stetigen Wandel nicht zu entsprechen. Naturwissenschaftliche Bildung kann somit nicht bei der Überlieferung isolierter Fakten stehen bleiben. Vielmehr muss es um den Erwerb einer generellen naturwissenschaftlichen Kompetenz in Sinne eines Anwendungsbezuges gehen (Bybee, 1997; Gross, 2006). Naturwissenschaftliche Kompetenz sollte Personen ermöglichen, lebenslang ungezwungen auf den Wegen der Erkenntnisgewinnung naturwissenschaftliche Phänomene zu verstehen und mit ihnen umzugehen, persönliche Entscheidungen treffen und an gesellschaftlichen Diskussionen teilnehmen zu können. Die referierten Ausführungen zeugen von einer funktionalen Auffassung naturwissenschaftlicher Kompetenz (Holbrook & Rannikmae, 2009). Diese spiegelt sich auch in der vorgelagerten Definition des allgemeinen Kompetenzbegriffs wider.

1.1 Definition des Begriffs fachübergreifende Kompetenz

Um die Annäherung an den Begriff naturwissenschaftliche Kompetenz und Kompetenz in Biologie voranzutreiben, wird zunächst eine trennscharfe Definition disziplinspezifischer Kompetenz benötigt. Der Begriff Kompetenz hat in der deutschen Bildungsforschung erst im Zuge des Paradigmenwechsels in den letzten 10 Jahren stark an Bedeutung gewonnen (Klieme & Leutner, 2006), erste Definitionsversuche gehen jedoch zurück bis in die siebziger Jahre des 20. Jahrhunderts. Unter der Vielzahl von Definitionen (Weinert, 2001) bleibt ein Großteil breit gefasst und kann nicht für eine Abgrenzung zu verwandten Begriffen wie Fähigkeiten oder Fertigkeiten herangezogen werden.

Erste Kompetenzdefinitionen sind in den Sprachwissenschaften angesiedelt; zunächst beschrieb Noam Chomsky Kompetenz als „kognitives Regelsystem, mit dessen Hilfe (sprachliche) [Klammersetzung v Verf.] Handlungen generiert werden können“ (Lenzen, 2001, S. 877). Kompetenz wird klar von deren Anwendung getrennt; die Anwendung hängt von weiteren Faktoren wie „außersprachliche Überzeugungen über den Sprecher und die Situation“ (Lenneberg, 1972, S. 485) ab. Chomskys Modell der linguistischen Kompetenz wurde zunehmend generalisiert und in der modernen Entwicklungspsychologie zur Charakterisierung der Ontogenese in z. B. numerischen oder räumlichen Gebieten verwendet (Weinert, 2001). In den folgenden Jahren beschrieben im deutschsprachigen Raum Lenneberg und Piaget den allgemeinen Kompetenzbegriff. Lenneberg arbeitete Kompetenz im Zusammenhang mit der Herausbildung von Sprache als grundlegende Fähigkeit von Menschen heraus, spezielle Fähigkeiten erwerben zu können (Lenneberg, 1972). Auch Piaget beschreibt eine idealisierte Kompetenzstruktur des Menschen (Lenzen, 2001). Diese allgemeinen Definitionsansätze mögen für erziehungswissenschaftliche, theoretische Diskurse ein geeignetes Fundament darstellen. Für eine empirische Untersuchung von Kompetenz im schulischen Rahmen sind sie jedoch zu breit gefasst. Daher werden im Folgenden Definitionen herangezogen, die stärker auf den Bildungsbereich ausgerichtet sind.

In Deutschland spielt in der bildungstheoretischen Diskussion der von Weinert (2001) formulierte Kompetenzbegriff eine herausgestellte Rolle. Demgemäß sind Kompetenzen „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2001, S. 27-28). Es handelt sich also um ein Leistungspotenzial im Sinne einer Disposition. Dieses Leistungspotential ermöglicht die Entwicklung von Fähigkeiten oder Fertigkeiten. Der Begriff Kompetenz skizziert somit Perspektiven von Entwicklungsmöglichkeiten eines Individuums. Auch diese Kompetenzdefinition bleibt sehr umfassend, da eine Vielzahl an unterschiedlichen Aspekten eingeschlossen wird. Sie dient

hauptsächlich der Abgrenzung der Kompetenz einer Person von seiner tatsächlichen Performanz in realen Situationen (Baartman, Bastiaens, Kirschner & van der Vleuten, 2007). Für eine empirische Erfassung disziplinspezifischer Kompetenz wird dagegen eine Fokussierung auf Teilaspekte dieser Definition als hilfreich angesehen (Klieme & Leutner, 2006; Weinert, 2001).

In der Weinertschen Definition wird Kompetenz in die Teilaspekte fachliche und fachübergreifende Kompetenz sowie Handlungskompetenzen gegliedert. Die Handlungskompetenzen beinhalten unter anderem motivationale und soziale Faktoren (Weinert, 2001). In einer differenzierteren Unterteilung werden Kompetenzen als intellektuelle Fähigkeiten, funktional auf bestimmte Situationen und kognitive Anforderungen bezogene Leistungsdispositionen, im Sinne motivationaler Orientierungen, Handlungskompetenzen und Metakompetenzen beschrieben (Klieme, Funke, Leutner, Reimann & Wirth, 2001). Im Schwerpunktprogramm Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen der Deutschen Forschungsgesellschaft (DFG) wird zur Ausschärfung des Kompetenzbegriffs daher nur ein Teilaspekt herausgegriffen. Kompetenzen werden als „kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Disziplinen beziehen“ (Klieme & Leutner, 2006, S. 4), definiert. Da die vorliegende Arbeit die kognitiven Aspekte naturwissenschaftlicher Kompetenz in den Blickpunkt nimmt, werden zur Definition von Kompetenz die kognitiven Aspekte der Weinertschen Definition herangezogen. Kompetenz wird in der vorliegenden Arbeit in Abwandlung bzw. Einschränkung des Weinertschen Kompetenzbegriffs und angelehnt an Klieme und Leutner (2006) wie folgt definiert:

Fachübergreifende Definition von Kompetenz

Kompetenz beinhaltet bei Individuen verfügbare bzw. erlernbare kognitive, kontextspezifische Leistungsdispositionen, die Individuen dazu befähigen, spezifische Probleme innerhalb einer akademischen Disziplin in variablen Situationen zu lösen.

Kompetenz ist demnach mehr als Wissen oder Erfahrungen (Weinert, 2001). Alle mentalen Ressourcen eines Individuums, die verwendet werden, um anspruchsvolle Aufgaben in verschiedenen inhaltlichen Disziplinen zu meistern, notwendiges deklaratives und prozedurales Wissen zu erlangen und gute Performanz zu erreichen spielen eine Rolle (Weinert, 2001). Diese mentalen Ressourcen werden von einigen Autoren auch als Wissensteile bezeichnet (Baartman et al., 2007).

Bereits bei dem fachübergreifenden Kompetenzbegriff wird der funktionale Aspekt von Kompetenz deutlich. Die Nutzung von Wissen und Fähigkeiten im Alltag und in der Gesellschaft steht im Vordergrund. Hierbei bleibt anzumerken, dass die spezifische Anwendung des Wissens von einer Disposition im Sinne einer inneren Bereitschaft aufgrund vorhandener kognitiver Ressourcen zu trennen ist. Diese Dispositionen sind erlernbar, jedoch nicht unmittelbar sichtbar. Sie schlagen sich vielmehr in den Handlungen und Entscheidungen einer Person nieder (Klieme, Hartig & Rauch, 2008).

Dieser erste Annäherungsversuch an eine Definition naturwissenschaftlicher Kompetenz bleibt auf einem allgemeinen Niveau stehen. Generelle Kompetenzen sind jedoch nur im Zusammenhang mit disziplinspezifischen Fähigkeiten und Fertigkeiten von Nutzen (Weinert, 2001). Die daraus resultierende Ausschärfung des Kompetenzbegriffs soll im Folgenden vorangetrieben werden. Die Mehrzahl der Definitionen liegt im aus deutscher Sicht interdisziplinären naturwissenschaftlichen Bereich vor. Hierbei handelt es sich um eine erhebliche Anzahl an Veröffentlichungen und Definitionen zu *scientific literacy* bzw. *science literacy* (z. B. Bybee, 1997; DeBoer, 2000; Gräber & Nentwig, 2002; Hodson, 1992; Hurd, 1998; Osborne, 2007; Shamos, 2002; Shen, 1975).

1.2 Beschreibung und Definition von naturwissenschaftlicher Kompetenz

Naturwissenschaftliche Kompetenz wird im fachdidaktischen Diskurs stark mit dem Konzept der *scientific literacy* bzw. *science literacy* in Verbindung gebracht. Grundlage der diversen *literacy* Konzeptionen ist eine Übertragung des linguistischen *literacy* Konzepts auf weitere akademische Disziplinen (z. B. in der Mathematik *mathematical literacy*). *Literacy* in seiner ursprünglichen Bedeutung beschreibt zunächst die Fähigkeit zu lesen und zu schreiben. Ohne diese beiden grundlegenden Fähigkeiten kann die Welt nur schwerlich erschlossen werden (Norris & Phillips, 2003). Erweiterte Bedeutungen des *literacy* Begriffs beinhalten auch Belesenheit, den adäquaten Umgang mit Texten im Alltag oder lernenden Umgang mit verschiedensten Textsorten. Es handelt sich also um eine Beherrschung von unterschiedlichen Prozessen in variablen, alltagsrelevanten und anwendungsbezogenen Situationen und der Möglichkeit erfahrungsbasiert zu lernen (Norris & Phillips, 2003), um die Beherrschung der Sprache der jeweiligen akademischen Disziplin auf einer abstrahierten Ebene (von Felden, 2003). Diese Prozesse sind notwendig, um kulturell signifikante Informationen zu interpretieren. Bei der Übertragung des *literacy* Konzepts auf weitere akademische Disziplinen wird der Versuch unternommen, disziplinspezifisch zu vermittelnde Kernkompetenzen – z. B. in den Naturwissenschaften oder in der Biologie – zu erfassen, die situationsadäquates Handeln ermöglichen. Dieser Übertragungsversuch wird auch in der Auffassung der AAAS zur Bedeutung und zum Nutzen von *scientific literacy* deutlich. In dem Projekt 2 061 der AAAS (1993) wird naturwissenschaftliche Kompetenz von Individuen wie folgt beschrieben:

“use the habits of mind and knowledge of science, mathematics and technology they have acquired to think about and make sense of many of the ideas, claims, and events that they encounter in everyday life. Accordingly, science literacy enhances the ability of a person to observe events perceptively, reflect on them thoughtfully, and comprehend explanations offered for them. In addition, those internal perceptions and reflections can provide the person with a basis for making decisions and taking action.” (AAAS, 1993, S. 322)

In Bybees Beschreibung (1997) von *scientific literacy* wird dieser Übertragungsversuch erneut deutlich:

„The phrase ‘scientific literacy for all learners’ expresses the major goal of science education – to attain society’s aspirations and advance individual development within the context of science and technology“ (Bybee, 1997, S. 69).

Das Konzept der *scientific literacy* im Sinne einer alltagsrelevanten naturwissenschaftlichen Kompetenz wurde von Shamos (1995) in Frage gestellt. Zwar negiert Shamos nicht den hohen Stellenwert, den naturwissenschaftliches Wissen (im weiteren Sinne) in der allgemeinen Öffentlichkeit haben sollte. Er hält jedoch insbesondere aufgrund der rasanten naturwissenschaftlichen Entwicklungen das Ziel einer *scientific literacy for all* für einen „*utopian dream*“ (Shamos, 1995, S. xii, s. a. Fischer, Klemm, Leutner, Sumfleth, Tiemann & Wirth, 2003). Außerdem benötigen nach seiner Auffassung die unterschiedlichen Zielgruppen (z. B.: zukünftige Naturwissenschaftlerinnen und Naturwissenschaftler, Interessensgruppen oder die breite Öffentlichkeit) unterschiedliches naturwissenschaftliches Wissen. Ein weiterer Kritikpunkt am Konzept *scientific literacy* betrifft die These, dass in der Schule erworbene Kompetenz im Erwachsenenalter beibehalten wird. Er vertritt daher die Meinung, dass naturwissenschaftliche Bildung eine *scientific awareness* zum Ziel haben sollte. Da *scientific awareness* zukünftigen Bürgerinnen und Bürgern noch nicht zu der Teilhabe an einer demokratischen und von den Naturwissenschaften stark beeinflussten Gesellschaft verhilft, sollte sie als eine erste Stufe auf dem Weg zu *scientific literacy* eingeordnet und nicht zum Ziel naturwissenschaftlicher Bildung werden. Die weiteren theoretischen Überlegungen zu *scientific literacy* des folgenden Unterkapitels geben einen differenzierteren Überblick über die vielfältigen Auffassungen zu diesem Konstrukt.

1.2.1 Theoretische Annäherungen an den Begriff *scientific literacy*

Unzählige internationale und nationale Publikationen unterstreichen die Zentralität des Begriffs *scientific literacy* im Zusammenhang mit dem Versuch der Verortung naturwissenschaftlicher Kompetenz (DeBoer, 2000; Gräber & Bolte, 1997; Gräber & Nentwig, 2002; Harlen, 2001; Hurd, 1998; Roberts & Gott, 2004; Shen, 1975). *Scientific literacy* steht dabei für den breiten und allumfassenden Zweck, dem naturwissenschaftliche Bildung dienen soll (Bybee, McCrae & Laurie, 2009). So liegt eine Vielzahl an Definitionsversuchen vor (z. B. DeBoer, 2000; Hurd, 1998; Lee, 1997; *National Research Council* [NRC], 2006; Shen 1975; Shortland 1988;), die das Konzept der *scientific literacy* aus unterschiedlichen Perspektiven beleuchten. Diese Definitionsversuche haben noch nicht zu einem Konsens über Inhalt und Struktur dieses Konstrukts geführt (DeBoer, 2000; Hodson, 1992; Keiser, Nelson, Norris & Szyszkiewicz, 1998; Osborne, 2007; O'Sullivan, Weiss & Askew, 1998). Einige Autoren sehen dem gegenüber das Konstrukt als umfassend und erschöpfend definiert an (Bybee, 2002; Deutsches PISA-Konsortium, 2002). DeBoer (2000) konnte jedoch in seinem Artikel über die Entwicklung der Definitionen neun verschiedene Ziele ausloten, die in der Rahmung von *scientific literacy* enthalten sind. Der Aspekt der Verknüpfung des Unterrichtens der Naturwissenschaften in der Schule mit den Alltagserfahrungen der Schülerinnen und Schüler liegt den meisten Konzeptualisierungen zu Grunde. Die Hauptunterschiede dieser Konzeptualisierungen liegen in dem Niveau, welches von einer naturwissenschaftlich literaten Person erwartet wird (Shamos, 1995).

Erstmalig wurde der Begriff *scientific literacy* von James Bryant Conant in dem von Watson und Cohen (1952) herausgegebenen Buch *General Education in Science* schriftlich erwähnt und 1958 von Hurd in das Lexikon der Naturwissenschaften aufgenommen (Bybee, 2002; Osborne, 2007); die folgende Entwicklung der theoretischen Auseinandersetzungen zur Ausschärfung der Begrifflichkeit kann in drei Phasen eingeteilt werden (Laugksch, 2000). Die erste Phase wird als Phase der Legitimation bezeichnet (ca. 1957-1963). Die zweite Phase der Interpretation (ca. 1974-1991) ist gefolgt von der Phase der weiteren Ausschärfung (seit 1983). Trotz dieser verhältnismäßig langen Entwicklungsphase der Konzeptualisierung von *scientific literacy* scheint eine Konsolidierung dieser Definitionen noch nicht gelungen. In der internationalen Diskussion ist kaum eine inhaltlich kohärente Rahmung zu erkennen. Vielmehr handelt es sich um ein diffuses, ungenau bestimmtes Konzept (Laugksch, 2000). Es folgt eine Vorstellung der in Deutschland bedeutsamen Definition von Bybee und unterschiedlicher Konzeptualisierungen internationaler und in Deutschland durchgeführter *large-scale assessment* Studien. Auf dieser Grundlage soll anschließend der Begriff naturwissenschaftliche Kompetenz definiert werden.

Bybee beschreibt das Hauptziel naturwissenschaftlicher Bildung als Entwicklung eines individuellen Verständnisses der Naturwissenschaften (und der Technologie). Als bedeutende

Aspekte naturwissenschaftlicher Literalität bezeichnet er das Vokabular, Konzepte, Prozesse und den Inhalt (Bybee, 1997). Er klassifiziert *scientific literacy* in vier hierarchisch aufeinander aufbauende Stufen naturwissenschaftlicher Grundbildung, die ein Kontinuum von passiven zu aktiven Dimensionen aufzeigen. Diese sollen kurz vorgestellt und am Beispiel der Biologie näher illustriert werden (Kattmann, 2003b). Da die prozedurale Stufe für die Theoriebildung der vorliegenden Arbeit von besonderer Bedeutung ist, wird diese – wie auch in der Veröffentlichung von Bybee – eingehender erläutert.

Nominale *scientific literacy* (Stufe 1) beinhaltet die Zuordnung von Fragestellungen, Themenstellungen und Begriffen zum naturwissenschaftlichen Bereich. Diese werden jedoch nicht im naturwissenschaftlichen Zusammenhang verstanden. Eine Schülerin bzw. ein Schüler auf dieser Stufe kann z. B. ein Laubblatt von einer Tannennadel unterscheiden.

Funktionale *scientific literacy* (Stufe 2) beinhaltet die Anwendung von Fachwissen in einem klar abgrenzbaren Rahmen, ohne dieses Wissen in weiteren Kontexten und in all seinen Facetten anwenden zu können. Das Fachwissen wird lediglich im naturwissenschaftlichen Sinn korrekt eingesetzt. Die Schülerin bzw. der Schüler auf dieser Stufe weiß, dass Blätter Teile eines Baumes mit speziellen Funktionen sind.

Das Erreichen konzeptueller und prozeduraler *scientific literacy* (Stufe 3) bringt ein Verständnis von zentralen naturwissenschaftlichen Ideen und Verfahren sowie deren Zusammenhänge innerhalb der Disziplin mit sich. Fundamentale naturwissenschaftliche Fähigkeiten und Konzepte dieser Stufe beinhalten:

- Fragestellungen und Konzepte identifizieren, die naturwissenschaftliche Untersuchungen bestimmen,
- naturwissenschaftliche Untersuchungen planen und durchführen,
- mittels Logik und Beweisführung naturwissenschaftliche Erklärungen und Modelle formulieren und überarbeiten,
- Erkennen und Analysieren von alternativen Erklärungen und Modellen und
- Kommunizieren und Verteidigen eines naturwissenschaftlichen Arguments.

Auf dieser Stufe hat die Schülerin bzw. der Schüler umfassende Konzepte verstanden und kann über (natur)wissenschaftliche Tätigkeiten zu weiterem Wissen auf diesem Gebiet gelangen. So wird z. B. das umfassende Konzept Ökosystem verstanden und durch das Beobachten des Verhaltens der unterschiedlichen Organismen eines Ökosystems das damit verbundene Besetzen von ökologischen

Nischen erkannt. Außerdem werden Zusammenhänge mit angrenzenden und übergeordneten Ökosystemen wahrgenommen.

Mehrdimensionale *scientific literacy* (Stufe 4) schließt das Verständnis der Besonderheiten naturwissenschaftlichen Denkens und die Fähigkeit zur Einordnung in soziale und kulturelle Zusammenhänge des persönlichen Lebens und in der Gesellschaft ein. Sie umfasst historische, philosophische und soziale Aspekte naturwissenschaftlicher Kompetenz (Bybee, 1997).

Nach Bybees Auffassung liegt innerhalb der naturwissenschaftlichen Kompetenz die Benennung und Verwendung auf einer niedrigeren Stufe, das konzeptuelle, prozedurale und mehrdimensionale Verständnis auf höheren Stufen. Diese Stufen naturwissenschaftlicher Kompetenz sind nicht fachspezifisch auf die drei Disziplinen Biologie, Physik und Chemie ausgerichtet. Eine empirische Prüfung dieser Konzeptualisierung für die einzelnen Disziplinen wie auch für die Naturwissenschaften steht noch aus. Dass diese noch nicht durchgeführt wurde, ist u. a. daran zu erkennen, dass Bybee von Stufen naturwissenschaftlicher Kompetenz ausgeht, während die *Science Functional Expert Group* (SFEG) der OECD diese eher als einzelne Dimensionen konzeptualisiert. Innerhalb von diesen Dimensionen können sich Individuen unterschiedlich entwickeln (Deutsches PISA-Konsortium, 2001). So werden die beschriebenen Stufen in die Analysen der PISA-Studien als voneinander zu unterscheidende Dimensionen aufgenommen (s. Seite 18).

Weitere Konzeptualisierungen verstehen naturwissenschaftliche Kompetenz als mehrdimensional und beinhalten den Wissens- und Prozessaspekt. So werden der *scientific literacy* z. B. die vier Elemente (a) konzeptuell, (b) kognitiv, (c) *Ideas-About-Science* und (d) sozial & affektiv zu Grunde gelegt (Osborne 2007). Mit diesen Elementen sind das konzeptuelle Wissen der Naturwissenschaften, die Fähigkeit zum schlussfolgernden Denken und wissenschaftlichen Denken, die essenziellen Bestandteile der Naturwissenschaften sowie die affektiven und sozialen Aspekte wie Motivation, Selbstkonzept oder das Klassenklima gemeint. Auch Hodson (1992) berücksichtigt die Mehrdimensionalität, indem er die sich gegenseitig beeinflussenden Hauptelemente (a) *learning science*, (b) *learning about science* und (c) *doing science* aufmacht. Während *learning science* die Aneignung und Entwicklung des konzeptuellen und theoretischen Wissens umfasst, beinhaltet das *learning about science* die Entwicklung eines Verständnisses der Charakteristiken und Methoden der Naturwissenschaften und die Wertschätzung der komplexen Interaktionen mit der Gesellschaft. *Doing science* ist die Vertiefung und Entwicklung von Expertise in *scientific inquiry* und Problemlöseprozessen. *Scientific inquiry* beinhaltet die wissenschaftlichen Erkenntnismethoden und schließt die Fähigkeiten ein, naturwissenschaftliche Fragestellungen und Konzepte zu identifizieren und naturwissenschaftliche Untersuchungen aufzubauen und durchzuführen (Bybee, 1997; Lee, 1997). Shavelson et al. (2008) skizzieren wie auch das NAGB (Winick et al., 2008) in Bezug auf die

kognitive Ebene von *science achievement* vier Wissenstypen bzw. Typen des schlussfolgernden Denkens innerhalb einer Disziplin und deren Zusammenspiel (s. Abbildung 1).

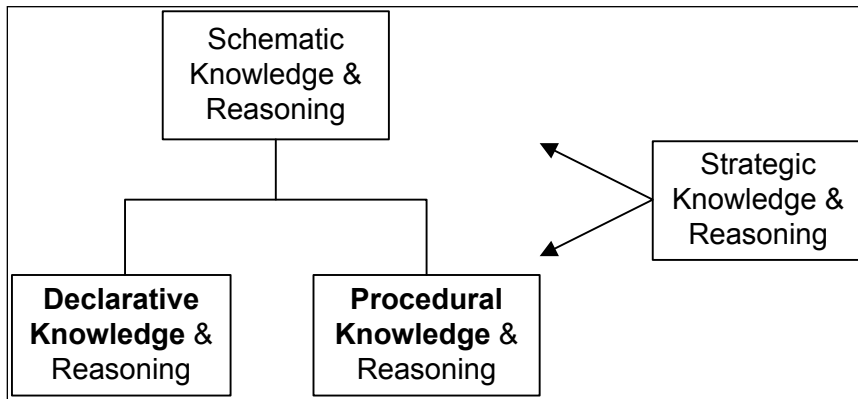


Abbildung 1: Entwurf der Wissens-/Reasoningtypen und ihrer Beziehungen zueinander

Aus „On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers“ von R. J. Shavelson, D. B. Young, C. C. Ayala, P. R. Brandon, E. M. Furtak, M. A. Riuз-Primo, 2008, *Applied Measurement in Education*, 21, S. 303. Copyright 2008 bei Richard J. Shavelson. Wiedergabe mit Genehmigung.

Das deklarative Wissen beinhaltet das *Wissen was* und somit das Fakten- und Konzeptwissen. Das prozedurale Wissen setzt sich sowohl aus dem Wissen über schrittweise Abfolgen als auch dem konditioniertes Wissen zusammen. Das schematische Wissen nimmt Bezug auf das Verbinden der erstgenannten Wissensarten über ein mentales Modell und das strategische Wissen ermöglicht die adressaten- und situationsgerechte Anwendung. Die Konzeptualisierung naturwissenschaftlicher Kompetenz von Gräber, Nentwig und Nicolson (2002) ist in Abbildung 2 dargestellt.

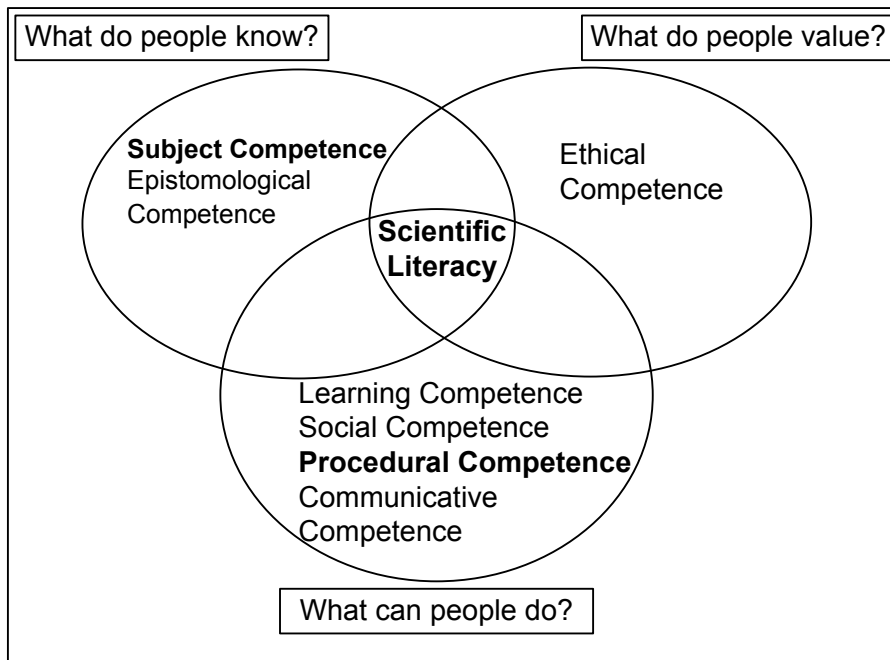


Abbildung 2: Kompetenzstrukturmodell naturwissenschaftlicher Kompetenz

Aus *Scientific Literacy – von der Theorie zur Praxis* (S. 137) von W. Gräber, P. Nentwig, P. Nicolson, 2002, in *Science Education – Past, Present, and Future* von Gräber, W.; Nentwig, P.; Koballa, T. & Evans, R., Opladen: Leske + Budrich. Copyright bei Springer Science and Business Media. Wiedergabe mit Genehmigung.

Beide Abbildungen machen deutlich, dass *scientific literacy* aus vielfältigen Kompetenzaspekten zusammengesetzt ist, wie z. B. ethische und soziale, kommunikative oder epistemologische Kompetenz bzw. schematisches, strategisches, deklaratives und prozedurales Wissen zusammengesetzt ist. Bei Gräber et al. (2002) wird naturwissenschaftliche Kompetenz in die drei Hauptkategorien: (a) was Personen wissen, (b) welche Werte Personen haben und (c) wie Personen handeln eingeteilt. Diese Unterteilung weist Ähnlichkeiten mit der Aufteilung von *scientific literacy* in die vier Komponenten: (a) intellektuelle, (b) gesellschaftliche und (c) interdisziplinäre Komponente sowie (d) die Komponente der Einstellungen auf (Holbrook & Rannikmae, 2009). Eine Behandlung all dieser Kompetenzaspekte des mehrdimensionalen Konstrukts (Hodson, 1992) ist im Rahmen der vorliegenden Arbeit weder möglich noch förderlich. Daher wird auf die in Abbildung 1 und Abbildung 2 fett hinterlegten Konzept- und Prozessaspekte von *scientific literacy* fokussiert. Die unterschiedlichen theoretischen Modelle zeigen, dass die strukturelle Zusammensetzung naturwissenschaftlicher Kompetenz noch nicht geklärt ist. Da die Bedeutung von Konzeptualisierungen internationaler *large-scale assessment* Studien in bildungstheoretischen Diskussionen stark gestiegen ist, werden auch diese für die Definition naturwissenschaftlicher Kompetenz herangezogen.

1.2.2 Konzeptualisierung und Operationalisierung naturwissenschaftlicher Kompetenz in *large-scale assessment* Studien sowie deren Ergebnisse

Bei der folgenden Darstellung der Konzeptualisierung und Operationalisierung von *large-scale assessment* Studien wird erneut deutlich, dass das Konzept- und Prozesswissen von den meisten Autorinnen und Autoren als integraler Bestandteil naturwissenschaftlicher Kompetenz hervorgehoben wird. Die aktuellen Ergebnisse von deutschen Schülerinnen und Schülern bezogen auf die Gesamtkompetenz sowie auf deren Teilkompetenzen zeigen differentielle Fähigkeitsmuster.

1.2.2.1 Konzeptualisierung naturwissenschaftlicher Kompetenz in *large-scale assessment* Studien

Einflussreiche Definitionen naturwissenschaftlicher Kompetenz im Zusammenhang mit *large-scale assessment* Studien gehen einerseits von Studien in den USA aus; andererseits geben internationale Studien zur Kompetenzmessung wichtige Impulse für deren Verortung. Das *National Assessment for Educational Progress* (NAEP) definiert im Zusammenhang mit Testungen in den Naturwissenschaften für die Veröffentlichungen zu den *The Nation's Report Card* wie folgt:

"Science proficiency depends upon the ability to know and integrate facts into larger constructs and to use the tools, procedures, and reasoning processes of science for an increased understanding of the natural world" (NAGB, 1995, zitiert nach O'Sullivan et al., 1998, S. 3).

Kompetenzmessungen in den Naturwissenschaften wurden und werden in Deutschland vor allem durch die PISA-Studien und die *Trends in International Mathematics and Science Study* (TIMSS) vorangetrieben.

In den unterschiedlichen PISA-Veröffentlichungen wird *scientific literacy* wie folgt definiert:

- PISA 2000 und PISA 2003

"the capacity to use scientific knowledge, to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity." (OECD, 2000, S. 76; OECD, 2003, S. 133).

▪ PISA 2006

“An individual’s:

- *Scientific knowledge and use of that knowledge to identify questions, to acquire new knowledge, to explain scientific phenomena, and to draw evidence-based conclusions about science-related issues.*
- *Understanding of the characteristic features of science as a form of human knowledge and enquiry.*
- *Awareness of how science and technology shape our material, intellectual, and cultural environments.*
- *Willingness to engage in science-related issues, and with the ideas of science, as a reflective citizen.”* (OECD, 2007, S. 34-35).

▪ PISA 2009

„requires an understanding of scientific concepts, as well as the ability to apply a scientific perspective and to think scientifically about evidence”. (OECD, 2010, S. 23).

Naturwissenschaftliche Kompetenz wird angelehnt an diese Definitionen in die gleichberechtigten Dimensionen Prozesse, Wissen und Verständnis sowie Anwendungskontext unterteilt. Da die PISA-Studie 2006 den Schwerpunkt Naturwissenschaften hatte, wird diese Untersuchung im Folgenden im Vordergrund stehen. In Tabelle 1 werden die ersten beiden Dimensionen der PISA-Studie 2006 zusammengefasst.

Tabelle 1: Inhaltliche Beschreibung der Dimensionen Prozesse und Wissen der naturwissenschaftlichen Kompetenz in der PISA-Studie 2006

Wissenskomponente	Prozesskomponente
Kennen naturwissenschaftlicher Fakten, Terme und Konzepte	Naturwissenschaftliche Fragestellungen erkennen
Entwicklung von Verständnis naturwissenschaftlicher Prinzipien, Gesetze, Methoden, Prozesse, Konzepte und Grenzen (<i>knowledge about science</i>)	naturwissenschaftliche Phänomene erklären
Fokussierung auf thematische Inhalte	Naturwissenschaftliche Evidenz nutzen Naturwissenschaftliches Forsuchen Naturwissenschaftliche Erklärungen

Basierend auf: OECD (2007) und PISA-Konsortium Deutschland (2007).

Anmerkungen. Ausführliche Beschreibung der Prozesskomponente siehe Tabelle 29 im Anhang.

Die Wissenskomponente wird über Themenbereiche aus den einzelnen naturwissenschaftlichen Disziplinen abgedeckt. Die Schülerinnen und Schüler sollen sowohl über Wissen über das naturwissenschaftliche Forsuchen als auch über naturwissenschaftliche Erklärungen verfügen. In der PISA-Studie 2006 waren die Themen für die Biologie Zellen, Menschen, Populationen, Ökosysteme und Lebensräume (OECD, 2007). Die Fokussierung auf dieses Fachwissen stand im naturwissenschaftlichen Unterricht in Deutschland bisher im Vordergrund (Duit et al., 2001; PISA-Konsortium Deutschland, 2007). In den USA konnte in den 80er Jahren gezeigt werden, dass die Schülerinnen und Schüler ein stark ausgeprägtes Fachwissen bei gleichzeitigem geringen Verständnis von den Naturwissenschaften aufweisen (Shamos, 1995). Ein Verständnis in beiden Wissensbereichen ist jedoch notwendig (Rost et al., 2004).

Die Prozesskomponente hingegen bildet die naturwissenschaftlichen Vorgehensweisen ab. Sie wird als unabhängig von der Wissensebene gesehen, kann jedoch nicht ohne Inhalte angewendet werden. Sie beinhaltet sowohl (a) naturwissenschaftliche Fragestellungen erkennen als auch (b) naturwissenschaftliche Phänomene erklären und (c) naturwissenschaftliche Evidenz nutzen. Hier scheinen im Besonderen Defizite bei den deutschen Schülerinnen und Schülern vorzuliegen (Baumert et al., 1997; Duit et al., 2001; Hammann, 2006; Klos, Henke, Kieren, Walpuski & Sumfleth, 2008; PISA-Konsortium Deutschland, 2007). So konnte gezeigt werden, dass im deutschen Physikunterricht das Anwenden von Wissen und das Schülerexperiment nur an dritter und vierter Stelle der Unterrichtsstrategien stehen (Baumert, 2002). Schon vor vier Jahrzehnten wurde kritisiert, dass beim Lehren nicht beide Kompetenzbereiche in Betracht gezogen werden (Layton, 1973). Auch wenn in den PISA-Studien der Jahre 2000, 2003 und 2009 die naturwissenschaftliche Kompetenz nicht im Vordergrund stand, wird auch in den Konzeptualisierungen dieser Erhebungswellen eine

Unterscheidung einer Wissens- und einer Prozesskomponente vorgenommen. So wurden in der PISA-Studie 2003 die Prozesse (a) Phänomene erklären, (b) wissenschaftliche Untersuchungen verstehen und (c) Schlussfolgerungen ziehen unterschieden (OECD, 2004), in der PISA-Studie 2009 (a) naturwissenschaftliche Fragestellungen erkennen, (b) naturwissenschaftliche Phänomene beschreiben, erklären und vorhersagen sowie (c) naturwissenschaftliche Evidenz nutzen, um Entscheidungen zu treffen (Klieme et al., 2010).

Bei dem dritten, quer zu den beiden Dimensionen liegenden Aspekt Anwendungskontext spielt der Einsatz des Wissens in unterschiedlichen alltagsnahen Situationen eine Rolle. Er beinhaltet die Fähigkeit, Probleme in Situationen der realen Welt zu lösen, welche uns als Individuum, als Mitglied einer lokalen Gemeinde und als Weltbürger beeinflussen können. In den letzten Jahren ist in Deutschland eine Hinwendung zu anwendungsorientiertem und kontextbezogenem Lernen zu verzeichnen (PISA-Konsortium Deutschland, 2007).

Die Konzeptualisierungen naturwissenschaftlicher Kompetenz der PISA-Studien sind an das NAEP angelehnt (Allen, Carlson & Zelenak, 1999; Winick et al., 2008). Im *Science Framework* vom NAEP von 1996 werden drei umfassende Themengebiete aufgespannt. Dies sind (a) die Themenfelder der Naturwissenschaften, (b) Wissen und Handeln in den Naturwissenschaften und (c) übergreifende Domänen. Die Themenfelder heißen Physik, Biologie und Geographie. Unter dem Oberbegriff Wissen und Handeln werden konzeptuelles Verständnis, naturwissenschaftliche Untersuchung und praktisches Schlussfolgern subsumiert. *Nature of science* und Themen in Form von *big ideas* sind schließlich die übergeordneten Domänen. Das *Science Framework* wurde für jüngste *large-scale assessment* Studien überarbeitet. Die nunmehr zwei Themengebiete lauten Inhalte der Naturwissenschaften und Handeln in den Naturwissenschaften (Winick et al., 2008). Die Inhalte sind in die drei Gebiete Physik und Biologie sowie Geographie und Weltraumwissenschaften unterteilt. Es wird weiterhin in Themen und Subthemen unterschieden. Für die Biologie werden die Konzepte Struktur und Funktion sowie Veränderungen von lebendigen Systemen behandelt. Das Handeln in den Naturwissenschaften besteht aus den Methoden Identifizierung und Verwendung von naturwissenschaftlichen Prinzipien sowie Verwendung von *scientific inquiry* und technologischen Designs (Winick et al., 2008).

Die Studien der *International Association for the Evaluation of Educational Achievement* (IEA), zu denen auch die TIMS-Studie zählt, beziehen sich auf die Taxonomie von Lernzielen nach Bloom (1974). Dieses umfassende, knappe Modell umfasst die Bereiche Kognition (Bloom, 1974), Affektivität (Krathwohl & Bloom, 1975) und Psychomotorik (nicht veröffentlicht) und dient der Analyse der Ergebnisse des Erziehungsprozesses. Die Taxonomie von Lernzielen im kognitiven Bereich stellt ein pädagogisches, logisches und psychologisches Klassifikationsschema dar und

schließt die Kategorien Wissen, Verstehen, Anwendung, Analyse, Synthese sowie Bewertung ein. Diese Kategorien sind hierarchisch zu verstehen; das Wissen ist z. B. Voraussetzung für das Verstehen. Die vier erstgenannten Kategorien finden sich in der Operationalisierung von TIMSS wieder. Eine Revision dieser Taxonomie von Anderson et al. (2001) unterscheidet zwischen den beiden Dimensionen Wissen und kognitive Prozesse. Die Wissensdimension besteht aus den Kategorien Fakten- und Konzeptwissen, prozeduralem Wissen sowie metakognitivem Wissen und dazugehörigen Fähigkeiten. Die kognitive Dimension besteht aus den Kategorien Erinnern, Verstehen, Anwenden, Analysieren, Bewerten und Kreieren. Auf Grundlage dieser Unterscheidung von kognitiven Lernzielen werden die internationalen TIMSS-Testaufgaben für die Jahrgangsstufe acht in die Bereiche (a) Faktenwissen, (b) konzeptuelles Verständnis sowie (c) Schlussfolgerungen und Analysen unterteilt.

In der ersten TIMS-Studie 1995 wurde *scientific literacy* als „Kommunikations- und Urteilsfähigkeit im naturwissenschaftlichen und technischen Bereich“ (Baumert et al., 1997, S. 61) definiert. Einzelne Dimensionen wurden für die Population II – also die Sekundarstufe I – nicht theoretisch herausgearbeitet. Die Operationalisierung naturwissenschaftlicher Kompetenz lässt jedoch auf eine Einbeziehung von fachlichen und erkenntnisgewinnungsbezogenen Komponenten schließen (s. Seite 21). So werden (a) naturwissenschaftliche Begriffe und Prinzipien sowie (b) naturwissenschaftliche Untersuchungsmethoden und Denkweisen unterschieden. Diese Bereiche sind wiederum in einzelne Aspekte differenziert worden. In der TIMS-Studie 2003 wurden die Naturwissenschaften in die Inhalts- und die kognitive Domäne unterteilt. Die kognitive Domäne gliedert sich in die drei Bereiche (a) Faktenwissen, (b) Konzeptverständnis sowie (c) Schlussfolgern und Analysieren (Beaton et al., 1996). *Scientific inquiry* wurde als übergreifende Domäne aufgenommen (Neidorf, Binkley & Stephens, 2006). Die Inhalte berücksichtigen neben den schulfachbezogenen Themenbereichen Biologie, Chemie, Physik und Geographie auch die Umweltwissenschaften. In der letzten TIMS-Studie 2007 wurde diese Konzeptualisierung geringfügig verändert. Die Umweltwissenschaften wurden nicht als ein eigenständiger Themenbereich definiert. Auf der kognitiven Seite wurden die drei Aspekte Wissen, Anwendung und Schlussfolgerungen herausgearbeitet (Mullis et al., 2005).

Es wird deutlich, dass sowohl autoren- als auch studienbezogen unterschiedliche Vorstellungen von *scientific literacy* vorliegen. Der Mehrzahl der Definitionen liegt jedoch die Idee von zwei grundlegenden Komponenten – der Wissens- und der Prozesskomponente – als einem immanenten Bestandteil von *scientific literacy* zu Grunde. Duit et al. (2001) fassten aus bereits vorliegenden Untersuchungen und theoretischen Arbeiten naturwissenschaftliche Wissensbereiche von *scientific literacy* zusammen. In ihrer Synthese wird deutlich, dass ein klarer Fokus auf den Wissens- und den

Prozessaspekt gelegt wird. Um dies zu verdeutlichen, werden in Tabelle 2 die verschiedenen Aspekte naturwissenschaftlicher Kompetenz der referierten Untersuchungen und Veröffentlichungen den Teilbereichen Konzept- und Prozesswissen von Kompetenz in den Naturwissenschaften zugeordnet.

Tabelle 2: Konzeptualisierung naturwissenschaftlicher Kompetenz der PISA-Studie 2006, der TIMS-Studie 1995/96, bei Duit et al. (2001) sowie bei Bybee (1997)

Konzeptualisierung	Konzeptwissen	Prozesswissen
Bybee (1997)	nominale Stufe funktionale Stufe konzeptuelle Stufe	prozedurale Stufe
PISA 2006 (PISA-Konsortium Deutschland, 2007)	Kennen naturwissenschaftlicher Fakten, Terme und Konzepte (<i>knowledge of science</i>) Verständnis naturwissenschaftlicher Prinzipien, Gesetze, Methoden, Prozesse, Konzepte und Grenzen (<i>knowledge about science</i>)	Naturwissenschaftliche Fragestellungen erkennen Naturwissenschaftliche Phänomene erklären Naturwissenschaftliche Evidenz nutzen
TIMSS 1995/96 (Baumert et al., 1997)	naturwissenschaftliche Begriffe und Prinzipien	Naturwissenschaftliche Untersuchungsmethoden und Denkweisen
Duit et al. (2001) Wissensbereiche	Naturwissenschaftliche Begriffe und Prinzipien	Naturwissenschaftliche Untersuchungsmethoden und Denkweisen Vorstellungen zur Natur der Naturwissenschaften Vorstellungen und Einstellungen zur Relevanz der Naturwissenschaften in der Gesellschaft und Technik
Duit et al. (2001) Kompetenzfacetten	Wissen von Fakten, Konventionen und Benennungen Wissen von Begriffen und Prinzipien Verstehen von Zusammenhängen zwischen Begriffen und Prinzipien Verstehen von Zusammenhängen im gesellschaftlichen Raum Anwenden von naturwissenschaftlichen Begriffen und Prinzipien in einer problemhaltigen Situation	Anwenden von naturwissenschaftlichen Untersuchungsmethoden und Denkweisen in einer problemhaltigen Situation Partizipation an Entscheidungen im gesellschaftlichen Raum sowie dem Wissen entsprechendes Handeln

Anmerkungen: PISA = Programme for International Student Assessment, TIMSS = Trends in International Mathematics and Science Study.

Es muss erwähnt werden, dass die Strukturierung von Kompetenz bzw. Wissen eine künstliche Unterteilung ist. In realen Situationen spielen die unterschiedlichen Dimensionen von Kompetenz bzw. Wissen zusammen (Hodson, 1992; OECD, 2000; Shavelson et al., 2008). Außerdem entsprechen der naturwissenschaftlichen Kompetenz weitere Handlungsaspekte wie die Bewertung, die der Partizipation an Entscheidungen im gesellschaftlichen Raum entspricht, oder Kommunizieren und Verteidigen eines naturwissenschaftlichen Arguments beinhaltet (KMK, 2005a). Diese werden in der vorliegenden Arbeit bewusst nicht näher betrachtet. Sie werden nur selten als eigenständige Dimensionen einer Kompetenz behandelt (OECD, 2000; O'Sullivan et al., 1998). In vorangegangenen Konzeptualisierungen wurden diese dimensional Aspekte des Öfteren mit der Prozesskomponente vermischt (Bybee, 1997; OECD, 2000). So werden Elemente der Kommunikationskompetenz in PISA 2006 unter einer globalen Prozesskomponente aufgeführt (s. Tabelle 1). Bezogen auf die theoretischen Ausführungen zu *scientific literacy* kann noch nicht von einer einheitlichen Auffassung von Dimensionen naturwissenschaftlicher Kompetenz ausgegangen werden. Um den Aufbau naturwissenschaftlicher Kompetenz bei Schülerinnen und Schülern zu ermöglichen, ist das Wissen um die Struktur von *scientific literacy* jedoch unabdingbar. Mit der vorliegenden Arbeit wird dies auf Grundlage eines umfangreichen Datensatzes möglich sein. Da nicht alle Aspekte naturwissenschaftlicher Kompetenz umfassend eingebunden werden können, sollen die beiden Kompetenzkomponenten Konzept- und Prozesswissen in der vorliegenden Arbeit den Fokus bilden und in der Definition im Vordergrund stehen. Die beiden Aspekte wurden ausgewählt, da auch in Theorien zu Wissenserwerb und -repräsentationen Evidenz für diese Unterscheidung enthalten ist (s. Seite 50). Daher wird naturwissenschaftliche Kompetenz in der vorliegenden Untersuchung wie folgt definiert:

Definition naturwissenschaftlicher Kompetenz

Naturwissenschaftliche Kompetenz beinhaltet bei Individuen verfügbare bzw. erlernbare kognitive, kontextspezifische Leistungsdispositionen, die sie dazu befähigen, naturwissenschaftliche Probleme in variablen Situationen zu lösen. Angelehnt an den Begriff der *scientific literacy* setzt sie sich aus einer Konzept- und Prozesskomponente zusammen. Die Konzeptkomponente beinhaltet hierbei das fachliche Wissen – z. B. über naturwissenschaftliche Konzepte, Begriffe, Prinzipien und Fakten – und das Verständnis über die Zusammenhänge zwischen ihnen. Die Prozesskomponente spiegelt den Anwendungsbezug wider und umfasst sowohl das Verständnis von den Naturwissenschaften als auch die erkenntnistheoretischen Fähigkeiten im Zusammenhang mit der Wissensgenerierung in den Naturwissenschaften.

Die erarbeitete Definition impliziert, dass naturwissenschaftliche Kompetenz aus den zwei Teilbereichen Konzept- und Prozesswissen besteht. Diese Unterteilung fügt sich in vorangegangene Konzeptualisierungen naturwissenschaftlicher Kompetenz (z. B.: DeBoer, 2000; Duit et al. 2001; Hodson, 1992; Lavonen & Krzywacki & Koistinen, 2012; OECD, 2000; O'Sullivan et al., 1998; Prenzel, Geiser, Langeheine & Lobemeier, 2003; Senkbeil, Rost, Carstensen & Walter, 2005). Da die verschiedenen Studien zu naturwissenschaftlicher Kompetenz diese Teilbereiche nicht nur unterschiedlich konzeptualisieren sondern auch unterschiedlich operationalisieren, werden diese in den folgenden beiden Kapiteln vorgestellt. Auch unterschiedliche Konzeptualisierungen und Operationalisierungen innerhalb der gleichen Studien über die Zeit zeugen von dem Wandel der Auffassung naturwissenschaftlicher Kompetenz.

1.2.2.2 Operationalisierung naturwissenschaftlicher Kompetenz in internationalen *large-scale assessment* Studien

In internationalen Schulleistungstudien wird eine umfassende naturwissenschaftliche Kompetenz getestet. Handlungsrelevantes Wissen sowie unterschiedliche Teilkompetenzen werden eingebunden. So wurden in der nationalen und internationalen PISA-Studie 2000 insgesamt Testaufgaben zum Faktenwissen (7), Graphikverständnis (11) und Verbalisieren von Sachverhalten (13) sowie zu mentalen Modellen (6) und Schlüsse ziehen (14) entwickelt (Deutsches PISA-Konsortium, 2001; Prenzel, Häußler, Rost & Senkbeil, 2002). In der PISA-Studie 2003 wurden insgesamt 34 Aufgaben zu naturwissenschaftlichen Prozessen erstellt. Davon behandelten 18 Aufgaben die Thematik Phänomene erklären, 7 wissenschaftliche Untersuchungen verstehen und 9 Schlussfolgerungen ziehen. Weitere 45 Aufgaben wurden zu Wissen über Naturwissenschaften entwickelt (PISA-Konsortium Deutschland, 2004).

In der PISA-Studie 2006 mit dem Schwerpunkt Naturwissenschaften wurden insgesamt 103 Testaufgaben zu naturwissenschaftlichen Thematiken entwickelt. Diese schließen auch die in Deutschland weniger gängigen Disziplinen Technologie, Astronomie und Geographie ein. Die Biologie wird über das Themengebiet „Lebende Systeme“ eingebunden, welches in die Aspekte Zellen, Menschen, Populationen, Ökosysteme und Biosphäre aufgegliedert ist. Die Besetzung der Teilbereiche von *scientific literacy* mit Testaufgaben der PISA-Studie 2006 sowie aus dem internationalen Test der TIMS-Studie 2007 wird in Tabelle 3 dargestellt.

Tabelle 3: Verteilung der Aufgaben der internationalen Naturwissenschaftstests PISA 2006 und TIMSS 2007 auf Teilbereiche naturwissenschaftlicher Kompetenz

TIMSS 2007		PISA 2006	
Teilbereiche	Testaufgaben	Teilbereiche	Testaufgaben
Wissen	84	Naturwissenschaftliche Fragestellungen erkennen	24
Anwenden	86	Naturwissenschaftliche Phänomene erklären	53
<i>Reasoning</i>	44	Naturwissenschaftliche Evidenz nutzen	31
Gesamt	214		108
Wissensbereiche		Wissensbereiche	
Biologie	76	Wissen über Naturwissenschaften (<i>scientific enquiry</i>)	25
Chemie	42	Wissen über Naturwissenschaften (<i>scientific explanations</i>)	21
Physik	55	Wissen über physikalische Systeme	17
Geographie	41	... lebende Systeme	25
		... Erd- Und Weltraumsysteme	12
		... technologische Systeme	8
Gesamt	214		108

Quellen: Martin, Mullis & Foy, 2008; OECD, 2007.

Anmerkungen. An der TIMSS-Studie 2007 nahmen in Deutschland nur die Schülerinnen und Schüler der Jahrgangsstufe vier teil. Die Auflistung umfasst die Aufgaben aus dem internationalen Test der Jahrgangsstufe acht.

PISA = *Programme for International Student Assessment*, TIMSS = *Trends in International Mathematics and Science Study*.

Für TIMSS 2007 wurden insgesamt 214 naturwissenschaftliche Aufgaben in drei Disziplinen entwickelt. Von diesen nimmt die Biologie mit 76 Aufgaben den größten Raum ein. Hiervon fallen 19 in den Themenbereich Verstehen einfacher Informationen, 13 Aufgaben decken das Verstehen komplexer Informationen ab, eine Aufgabe betrifft das Konzeptualisieren und Problemlösen und schließlich acht Aufgaben das Experimentieren und die Beherrschung von Verfahren. Bezogen auf den Inhalt wurden wiederum Aufgaben zu 14 Themengebieten wie z. B. (a) Merkmale und Klassifikation von Lebewesen oder (b) Photosynthese und Atmung und (c) Lebenszyklen von Organismen entwickelt (Martin, Mullis & Foy, 2008). Für die internationale Berichterstattung wurden die Ergebnisse auf einer naturwissenschaftlichen Gesamtskala berichtet. In der nationalen

Erweiterung wird eine Darstellung getrennt nach den Disziplinen Biologie, Chemie und Physik vorgenommen, nach weiteren Untertests wird nicht differenziert (Baumert et al., 1997).

Im Zusammenhang mit den PISA-Studien 2000, 2003 und 2006 wurden nationale Ergänzungsstudien – sogenannte PISA-E Studien – durchgeführt, die auch zur Weiterentwicklung der Tests führten. In der PISA-Studie 2003 wurde sowohl eine nationale Erweiterung als auch eine Längsschnittstudie mit zwei Messzeitpunkten durchgeführt. Für die Längsschnittstudie wurde eine Teilmenge der PISA-E-Aufgaben verwendet. Die nationalen Erweiterungen spiegeln in einem höheren Maße als die internationalen Studien die Curricula der Bundesländer wider. Neben dieser curricularen Ausrichtung werden in der PISA-E Studie 2003 die kognitiven Prozesse (a) Bewerten, (b) divergentes Denken, (c) Graphikverständnis, (d) konvergentes Denken, (e) mentale Modelle, (f) Umgang mit Zahlen und (g) Sachverhalte verbalisieren konkretisiert. Jeweils eine Aufgabe zu einem der neun naturwissenschaftlichen Themen bestand aus sieben Teilaufgaben. Die sieben Teilaufgaben decken innerhalb jeder Aufgabe die sieben kognitiven Teilkompetenzen ab. Somit besteht die nationale Erweiterung der PISA-Studie 2003 aus 63 dieser Teilaufgaben. Im Jahr 2006 wurde die PISA-E Studie um Aufgaben zu den Bildungsstandards erweitert. Seit 2009 bestehen sie ausschließlich aus Aufgaben zu den Bildungsstandards. Im ersten Durchgang wurden die Mathematik und die Sprachen näher untersucht. Die PISA-Studie 2012 wird momentan mit den Aufgaben zu den Bildungsstandards in den Fächern Biologie, Chemie und Physik durchgeführt.

Im Rahmen des NAEP wurden 2009 und 2011 in den USA die zweijährlich stattfindenden Untersuchungen zu *The Nation's Report Card* durchgeführt, in denen auch Kompetenz in den Naturwissenschaften berücksichtigt wird (*National Center for Education Statistics* [NCES], 2011, NCES, 2012). Die Erhebungen finden in den Jahrgangsstufen 4, 8 und 12 statt. In der Jahrgangsstufe acht waren 2009 30 % der entwickelten Aufgaben auf dem Gebiet *life science* angesiedelt. Dieses Themengebiet war wiederum in (a) Strukturen und Funktionen von lebenden Systemen und (b) Veränderungen in lebenden Systemen differenziert. Die Prozessdimension besteht zu 25 % aus Identifizieren naturwissenschaftlicher Prinzipien, zu 35 % aus Verwendung naturwissenschaftlicher Prinzipien, zu 30 % aus Verwenden von *scientific inquiry* und zu 10 % aus technologischen Designs. Unter der Prozessdimension wurden des Weiteren die vier kognitiven Anforderungen (a) Wissen was, (b) Wissen wie, (c) Wissen warum sowie (d) Wissen, wann und wo Wissen angewendet werden sollte, unterteilt (Winick et al., 2008). Diese Unterteilung ist an das Modell von Shavelson et al. (2008) angebunden. Zur Testung der Prozessdimensionen bearbeitete eine Teilstichprobe der Schülerinnen und Schüler neben einem *paper-and-pencil* Test zusätzlich *hands-on* Aufgaben sowie Computeraufgaben. In dieser Teilstichprobe wurde die Testung um 30 Minuten verlängert (Winick et al., 2008). Die Operationalisierung der Themengebiete und der Prozessdimension von Kompetenz in

den Naturwissenschaften veränderte sich zur Erhebung 2011 nicht. In der Stichprobe der Jahrgangsstufe acht wurden keine praktischen Testungen oder alternative Computertestungen durchgeführt (NCES, 2012).

An dieser knappen Zusammenfassung zur Operationalisierung naturwissenschaftlicher Kompetenz wird deutlich, dass in den internationalen und nationalen Untersuchungen unterschiedliche Teilkompetenzen und Themengebiete Anwendung finden. Die Anzahl der eingesetzten Aufgaben pro Teilkompetenz reicht jedoch selten aus, um zu einer belastbaren Berichterstattung auf dieser Detailebene zu gelangen. Einzig in den PISA-Studien wurden erste Möglichkeiten einer differenzierten Berichterstattung genutzt (Deutsches PISA-Konsortium, 2001; PISA-Konsortium Deutschland, 2004; PISA-Konsortium Deutschland, 2007). Für die Berichterstattung weiterer internationaler und nationaler Untersuchungen werden größtenteils übergreifende Mittelwertvergleiche herangezogen. Die vorliegende Arbeit wird dieses Defizit angehen, indem Analysen zu den Teilbereichen Konzept- und Prozesswissen für die Disziplin Biologie durchgeführt werden.

Wie mehrfach gezeigt werden konnte, besteht die naturwissenschaftliche Kompetenz aus verschiedenen Facetten, von denen in dieser Arbeit das Konzept- und Prozesswissen untersucht werden sollen. Für diese beiden Kompetenzteilbereiche liegen vereinzelt Ergebnisse zu deren Zusammenhang und Unterscheidung in unterschiedlichen Disziplinen vor. Für die Naturwissenschaften und deren Teildisziplinen Biologie, Chemie und Physik ist jedoch noch nicht zufriedenstellend geklärt, ob und welche getrennten Dimensionen (Deutsches PISA-Konsortium, 2001; Klos et al., 2008) oder hierarchischen Bestandteile (Bybee, 1997) angenommen werden können. Die wenigen bereits bestehenden Untersuchungen werden im folgenden Kapitel behandelt.

1.2.2.3 Ergebnisse zur Struktur naturwissenschaftlicher Kompetenz

Im Zusammenhang mit den Konzeptualisierungen von *scientific literacy* wurden Vermutungen angestellt, welcher Prozentsatz der US-amerikanischen Bevölkerung nicht als *scientific literate* bezeichnet werden kann. In der Kritik von Shamos (1995) wurden weit über 90 % der Bevölkerung als nicht *scientific literate* titulierte, von dem *International Center for the Advancement of Scientific Literacy* (ICASL) 93 % (Westby & Torres-Velásquez, 2000). In der vorliegenden Arbeit wird zunächst die naturwissenschaftliche Kompetenz der deutschen Schülerinnen und Schüler in ihrer Gesamtheit kurz umrissen. Hierzu liegen Ergebnisse aus den drei Untersuchungszyklen der PISA und der TIMS-Studie vor.

In der PISA-Studie 2006 liegt die naturwissenschaftliche Gesamtkompetenz der deutschen Schülerschaft mit einem Punktwert von 516 Punkten ($SD = 100$ Punkte), im Jahr 2009 mit 520 Punkten ($SD = 101$ Punkte) statistisch signifikant über dem OECD-Durchschnitt von 500 Punkten

(Klieme et al., 2010; PISA-Konsortium Deutschland, 2008). Im Jahr 2003 hingegen erlangten die Schülerinnen und Schüler mit nur 502 Punkten den OECD-Mittelwert (PISA-Konsortium Deutschland, 2004), 2000 lagen sie mit 487 Punkten unter dem OECD-Durchschnitt (Deutsches PISA Konsortium, 2001). Der Trend weist also auf einen Anstieg der naturwissenschaftlichen Bildung deutscher Schülerinnen und Schüler hin. Die berichtete Effektstärke von $d = 0.17^1$ (Hammann & Prenzel, 2008) fällt in Anbetracht der Tatsache, dass es sich um eine Mehrkohorten-Untersuchung über die Zeit handelt, praktisch relevant aus. Es wird jedoch angemerkt, dass trotz dieser Verbesserung in den naturwissenschaftlichen Leistungen der Schülerinnen und Schüler, der Anteil im unteren Leistungsbereich noch zu hoch ausfällt (PISA-Konsortium Deutschland, 2007). In der stärker curricular ausgerichteten TIMS-Studie erreichten die Schülerinnen und Schüler der Jahrgangsstufe acht 531 Punkte ($M_{\text{international}} = 516$ Punkte, $\text{Max}_{\text{international}} = 607$ Punkte, $\text{Min}_{\text{international}} = 326$ Punkte) (Beaton et al., 1996). Im Fach Biologie erzielten die deutschen Schülerinnen und Schüler mit 63 % richtig gelösten Aufgaben das beste disziplinbezogene Ergebnis. Die folgenden Ergebnisse zu naturwissenschaftlichen Teilkompetenzen aus den nationalen Tests geben einen detaillierten Einblick in die Kompetenzen der Schülerinnen und Schüler der Sekundarstufe I.

In den PISA-Erhebungen der Jahre 2000 und 2003 wurden für Teilkompetenzen naturwissenschaftlicher Kompetenz je nach Erhebung Mittelwerte und Standardabweichungen der Aufgabenschwierigkeiten bzw. Schwierigkeitsparameter auf einer Logit-Skala berichtet (s. Seite 120). Die mittleren Schwierigkeitsparameter der Teilkompetenzen der PISA-Erhebung 2000 betragen:

- Faktenwissen ($p = 1.59$),
- Mentale Modelle ($p = 1.16$),
- Sachverhalte verbalisieren ($p = 0.89$),
- Schlüsse ziehen ($p = 0.45$) und
- Graphikverständnis ($p = -0.35$)

(Deutsches PISA-Konsortium, 2001).

¹ Effektstärken sind ein Maß für die praktische Bedeutsamkeit eines Unterschiedes zwischen zwei Werten. Je höher die Effektstärke desto praktischer relevanter der gefundene Unterschied.

Die Mittelwerte der Aufgabenschwierigkeiten der Teilkompetenzen der nationalen PISA-Erhebung 2003 lauten:

- Mentale Modelle ($M = 1.76$, $SD = 1.43$),
- Konvergentes Denken ($M = 1.25$, $SD = 1.35$),
- Umgang mit Zahlen ($M = 0.82$, $SD = 0.70$),
- Umgang mit Graphiken ($M = 0.51$, $SD = 0.64$),
- Divergentes Denken ($M = 0.15$, $SD = 0.72$),
- Bewerten ($M = -0.17$, $SD = 0.80$) und
- Sachverhalten verbalisieren ($M = -0.21$, $SD = 0.74$)

(PISA-Konsortium Deutschland, 2004).

Der Umgang mit graphischen Repräsentationen und divergentem Denken fällt den Schülerinnen und Schülern offenbar besonders schwer. Da die Naturwissenschaften in der Erhebung 2006 im Fokus standen, konnten erstmals detailliertere Untersuchungen zu Teilkompetenzen durchgeführt werden. Wie auch in der Gesamtkompetenz liegen die deutschen Schülerinnen und Schüler über dem OECD-Durchschnitt, der jeweils 500 Punkte beträgt. Der niedrigste Wert liegt bei der Skala naturwissenschaftliche Fragestellungen erkennen ($M = 510$ Punkte, $SD = 98$); im mittleren Bereich befindet sich die Skala naturwissenschaftliche Evidenz nutzen ($M = 515$ Punkte, $SD = 115$). Diese beiden Teilkompetenzen sind nach (Mayer, 2007) im Prozesswissen naturwissenschaftlicher Kompetenz anzusiedeln. Der höchste Wert liegt bei der Skala naturwissenschaftliche Phänomene erklären ($M = 519$, $SD = 103$), einer dem Konzeptwissen zugehörigen Teilkompetenz. Die Werte liegen nah beieinander, die Schwankungen zwischen den Teilbereichen fallen in Deutschland jedoch höher aus als die Schwankungen der OECD-Mittelwerte der Teilkompetenzen (PISA-Konsortium Deutschland, 2007). Für diese Teilkompetenzen sind unterschiedliche Leistungsprofile für Jungen und Mädchen zu verzeichnen. So sind die Mädchen besser im Erkennen naturwissenschaftlicher Fragestellungen ($M_{\text{Mädchen}} = 518$ Punkte vs. $M_{\text{Jungen}} = 502$ Punkte²) während die Jungen im naturwissenschaftliche Phänomene erklären ($M_{\text{Mädchen}} = 508$ Punkte vs. $M_{\text{Jungen}} = 529$ Punkte) besser abschneiden. Für den Teilbereich naturwissenschaftliche Evidenz nutzen liegen mit $M_{\text{Mädchen}} = 513$ Punkten und $M_{\text{Jungen}} = 517$ Punkten beide Gruppen ungefähr gleich auf (PISA-Konsortium Deutschland, 2008).

Vertiefende Unterrichtsanalysen auf Grundlage der Schülerdaten der PISA-Studie 2006 zeigten, dass im deutschen Naturwissenschaftsunterricht zwar zahlreiche interaktive Elemente wie *hands-on*

² In der Veröffentlichung wurden die Standardabweichungen für die Mittelwerte der Jungen und der Mädchen nicht mit abgedruckt.

Aktivitäten eingesetzt und Experimente im naturwissenschaftlichen Unterricht durchgeführt werden. Dennoch liegt der Schwerpunkt auf angeleiteten Aktivitäten, sodass nur selten der gesamte Prozess des naturwissenschaftlichen Denkens von den Schülerinnen und Schülern nachvollzogen wird (Fischer, Klemm, Leutner, Sumfleth, Tiemann & Wirth, 2005). Das Entwickeln eigener Fragestellungen und deren Testung in naturwissenschaftlichen Versuchen werden nur selten praktiziert (Keiser et al., 1998; PISA-Konsortium Deutschland, 2007). Vielmehr ist für den naturwissenschaftlichen Unterricht typisch, auf eine Variable zu fokussieren, die im Einklang mit einer bestehenden oder zu entwickelnden Theorie den Unterschied ausmacht (Zimmerman, 2007). Der im naturwissenschaftlichen Unterricht weit verbreitete bestätigende Versuch unterscheidet sich deutlich vom Forschungsversuch (Eschenhagen, Kattmann & Rodi, 1998). So konnte gezeigt werden, dass zwei Typen von Unterricht vorherrschend sind: lehrerzentrierter Unterricht mit Demonstrationsexperimenten und lehrerzentrierter Unterricht mit Anteilen von Schülerexperimenten (Baumert, 2002; Fischer et al., 2003; Seidel, Prenzel, Wittwer & Schwindt, 2007). Die Schülerexperimente können in imitatorische bzw. organisatorische und konzeptuelle Experimente unterteilt werden (Fischer et al., 2003). Diese Zweiteilung wird auch als *hands-on* versus *minds-on* diskutiert (Hammann, Phan & Bayrhuber, 2007). In dem internationalen Vergleich zu Zusammenhängen von Unterrichtsmustern mit den Kompetenzwerten und dem Interesse der Schülerinnen und Schüler zeigte sich, dass förderlicher Unterricht „einerseits Gelegenheiten zur selbstständigen Entwicklung und Durchführung von Experimenten bietet, andererseits Wert ... auf kognitiv aktivierende Tätigkeiten wie das Schlussfolgern und das Generieren von Ideen“ (PISA-Konsortium Deutschland, 2007, S. 174) legt.

Rost et al. (2004) führten weiterführende Untersuchungen mit den Daten des nationalen und internationalen Tests der PISA-Studie 2003 durch. Zwischen dem eher curricular ausgerichteten nationalen Test, der aber auch die kognitiven Kompetenzen näher untersuchte, und dem internationalen Test konnten keine Unterschiede in den Leistungen ermittelt werden (PISA-Konsortium Deutschland, 2005). Dennoch konnten die Jungen in den kognitiven Teilkompetenzen Faktenwissen und Umgang mit mentalen Modellen bessere Leistungen als die Mädchen erzielen. Die Unterschiede liegen bei 11.43 und 10.46 Kompetenzpunkten. Diese Geschlechterunterschiede nehmen für die Teilkompetenzen Sachverhalt verbalisieren (7.83) und Graphikverständnis (1.16) ab. Bei der Teilkompetenz Schlüsse ziehen kehrt sich das Verhältnis um (-0.68) (Rost et al., 2004). Auch die Unterschiede zwischen den Bundesländern sind nicht primär im Niveau, sondern vielmehr in den Leistungsprofilen zu suchen. Trotz eines tendenziellen Nord-Süd-Gefälles in den Mittelwerten sind differenzierte Befunde für das Faktenwissen und die weiteren kognitiven Kompetenzen zu erkennen. Die neuen Bundesländer zeigen höhere Werte im Faktenwissen. Eine weitere Untersuchung zu Unterschieden in den beiden Teilkompetenzen Prozess- und Konzeptwissen zeigte einen höheren

Zusammenhang des Prozesswissens mit dem sozioökonomischen Hintergrund (Turmo, 2004). Es wird deutlich, dass Informationen aus Untersuchungen von Teilkompetenzen detaillierte Förderung von schulischem Lernen ermöglichen.

Im Zuge der Feldstudie des nationalen Naturwissenschaftstests zur PISA-Studie 2003 wurde von Senkbeil et al. (2005) eine Veröffentlichung zu den Facetten naturwissenschaftlicher Kompetenz vorgelegt. Gezielt wurden 70 Aufgaben entwickelt, die in den drei Disziplinen Biologie, Physik und Chemie 10 Basiskonzepte (z. B. Atmung und Fotosynthese) und die sieben kognitiven Anforderungen (a) konvergentes und (b) divergentes Denken, (c) Umgang mit Graphen, (d) mit mentalen Modellen und (e) mit Zahlen, (f) Sachverhalte verbalisieren sowie (g) Bewertung/Evaluation) abbilden. Die kognitiven Anforderungen sind allgemeine mentale Prozesse, nicht disziplinspezifische Kompetenzen. Diese wurden an 1 955 15-jährigen Schülerinnen und Schülern in vier Bundesländern getestet. Es wurden verschiedene Mehrfacettenmodelle verglichen, die unterschiedliche Dimensionen spezifizieren. Ein Siebenfacettenmodell entsprach am besten der Struktur der empirischen Daten. Dieses definiert sieben kognitive Komponenten und keine Inhaltsdimension. Die Korrelationen dieser kognitiven Anforderungen untereinander liegen zwischen $r = .49$ (mentale Modelle und Bewerten) und $r = .89$ (Sachverhalte verbalisieren und Bewerten).

Neben den deskriptiven Befunden zu den Schülerleistungen in den Naturwissenschaften und ihren Teilkompetenzen wurden einige Analysen zu den Zusammenhängen letzterer berichtet. In der PISA-Studie 2000 zeigte sich, dass der naturwissenschaftliche Teilbereich Faktenwissen mit den eher kognitiven Aspekten naturwissenschaftlicher Kompetenz mit Korrelationen von $r = .58$ bis $r = .75$ in einem starken Zusammenhang steht (Deutsches PISA-Konsortium, 2003). Im Rahmen der Erhebung 2006 konnten diese Zusammenhänge noch detaillierter untersucht werden. In Tabelle 4 werden die Interkorrelationen der naturwissenschaftlichen Teilbereiche dargestellt.

Tabelle 4: Messfehlerbereinigte Korrelationen zwischen den naturwissenschaftlichen Teilkompetenzen der PISA-Studie 2006 (OECD, 2009)

	<i>Identifying scientific issues</i>	<i>Using scientific evidence</i>
<i>Identifying scientific issues</i>		.91 (.90)
<i>Explaining phenomena scientifically</i>	.90 (.89)	.93 (.90)
<i>Using scientific evidence</i>	.90 (.89)	.93 (.93)

Quelle: OECD, 2009

Anmerkungen. Die Angaben entsprechen den Korrelationen für die OECD-Länder, jeweils in Klammern stehen die Korrelationen für alle teilnehmenden Länder.

Es wird ersichtlich, dass die kognitiven Teilkompetenzen naturwissenschaftlicher Kompetenz stark miteinander korrelieren. Ähnlich hohe Korrelationen der Teilkompetenzen der Naturwissenschaften zwischen $r = .61$ in der TIMSS-Untersuchung und $r = .85$ in den PISA Untersuchungen wurden auch in einer jüngst veröffentlichten PISA-Kritik gezeigt (Rindermann, 2006). In einer Studie des *Berkeley Evaluation and Assessment Research Institutes* (BEAR) wurden die Dimensionen *Evidence and Tradeoffs* (ET), *Designing and Conducting Investigations* (DCI), *Understanding Scientific Concepts* (UC) und *Communicating Scientific Information* (CSI) näher untersucht (Draney & Peres, 1998). Hierfür nahmen 700 Schülerinnen und Schüler an einer Längsschnittstudie teil, die an einem speziellen einjährigen Curriculum für die Naturwissenschaften angekoppelt war. Die vierdimensionalen Modelle passten zu allen Messzeitpunkten besser auf die Daten als die eindimensionalen Modelle. Die Korrelationen zwischen den Dimensionen DCI – kongruent zum Prozesswissen – und UC – kongruent zum Konzeptwissen – lagen zwischen $r = .56$ und $.77^3$.

Über die Konzeptualisierungen internationaler Schulleistungstudien hinaus liegen kaum empirische Untersuchungen zur Dimensionalität naturwissenschaftlicher Kompetenz vor. Im deutschsprachigen Raum wird einzig in der Untersuchung von Klos et al. (2008) der Zusammenhang zwischen den Teilbereichen Fachwissen und naturwissenschaftliche Arbeitsweisen untersucht. Ziel der Studie ist die Validierung eines Tests zum Prozesswissen in der Chemie. Der Test zu naturwissenschaftlichen Arbeitsweisen (NAW-Test) besteht aus 39 Aufgaben, die den Inhaltsbereichen Idee/Hypothese (15 Testaufgaben), experimentelle Umsetzung (14 Testaufgaben) und experimentelle Schlussfolgerung (10 Testaufgaben) zuzuordnen sind. Der Test zielt also eher auf naturwissenschaftsbezogenes Prozesswissen im Sinne der Erkenntnisgewinnung als auf das Problemlösen ab (s. Seite 61). Der zur Validierung eingesetzte Fachwissenstest beinhaltet 66 Aufgaben, deren Inhalte und Konzeptualisierung nicht explizit beschrieben werden. Die Ausführungen lassen vermuten, dass es sich eher um einen Vorwissenstest als um einen Test zum konzeptuellen Verständnis handelt. In einer ersten Erhebung wurden 600 Schülerinnen und Schüler der siebten Klassenstufe getestet. Für 139 Schülerinnen und Schüler der Jahrgangsstufe 12 wurde der Test modifiziert, um ihn an die Voraussetzungen der Population anzupassen. Der Fachwissens- und der NAW-Test wiesen bei den jüngeren Schülerinnen und Schülern mit einer nicht statistisch signifikanten Korrelation von $r = .07$ nur einen sehr schwachen Zusammenhang auf. Dies wird von den Autoren mit dem fehlenden Fachwissen von Schülerinnen und Schülern dieser Klassenstufe erklärt. Dahingegen fällt der Zusammenhang in der Stichprobe der älteren Schülerinnen und Schüler mit $r = .68$ sehr stark aus. Um die Zusammenhänge näher zu untersuchen, wurden in einer zusammengefassten Stichprobe aus Elft- und Zwölftklässlern ($N = 218$) Strukturgleichungsmodelle

³ Eine Korrelation lag bei $r = .19$. Diese sollte nach Aussagen der Erstautorin nicht interpretiert werden, da die Reliabilität der Skala zu gering ausfällt. (K. Draney, pers. Mitteilung, 16. 02. 2011)

zum Zusammenhang der beiden Tests gerechnet. In einem ersten Modell wurden zwei latente Konstrukte – das Fachwissen und naturwissenschaftliche Arbeitsweisen – angenommen. Die Korrelation zwischen diesen beiden latenten Konstrukten fiel so hoch aus ($r = .85$), dass ein zweites genestetes Modell aufgestellt wurde, in welchem die Korrelation der beiden Konstrukte auf Eins gesetzt wurde. Der Vergleich der Fit-Indizes beider Modelle (Modell 1: Bentler *Comparative Fit Index* [Bentler CFI] = 1, *Root Mean Square of Approximation* [RMSEA] = .00 und *Standardized Root Mean Square Residual* [SRMR] = .02; Modell 2: Bentler CFI = .94, RMSEA = .07 und SRMR = .08) zeigt, dass das zweidimensionale Modell besser auf die empirischen Daten passt. Die Fit-Indizes werden im methodischen Kapitel erläutert.

Zusammenfassend kann festgestellt werden, dass die Subskalen naturwissenschaftlicher Kompetenz unterschiedlicher Studien stark miteinander zusammenhängen. Für die Kompetenz in Biologie liegen keine umfassenden theoretischen Modelle oder empirischen Untersuchungen zu deren Struktur vor. Die Unterteilung in verschiedene Aspekte wird zunehmend in die Schulcurricula der Bundesländer, wie z. B. dem Rahmenlehrplan für die Sekundarstufe I für das Fach Biologie des Landes Berlin (Senatsverwaltung für Bildung, Jugend und Sport [SBJs], 2006), aufgenommen. Daher werden im folgenden Kapitel insbesondere bildungspolitische Dokumente zur Verortung von Kompetenz in Biologie herangezogen.

1.3 Beschreibung und Definition von Kompetenz in Biologie

Im deutschen Kontext werden die Naturwissenschaften in der Primarstufe im Rahmen des Sachunterrichts interdisziplinär (z. B.: Ministerium für Bildung, Jugend und Sport des Landes Brandenburg [MBJS], SBJs & Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg-Vorpommern, 2004) in den Sekundarstufen überwiegend in den getrennten Disziplinen Biologie, Chemie und Physik unterrichtet. Die angegliederten Fachdidaktiken und deren Forschungsbestrebungen sind ebenfalls disziplinär angelegt. In den letzten Jahren wurden vermehrt Projekte zur Zusammenführung der Fachdidaktiken initialisiert. So startete z. B. im Herbst 2010 das Humboldt ProMINT-Kolleg an der Humboldt-Universität Berlin. In diesem Projekt werden disziplinübergreifende Strategien und Unterrichtskonzepte für die MINT-Fächer (Naturwissenschaften, Informatik, Technik und Mathematik) entwickelt (Schön, Kramer & Buttner, 2010).

Empirische, repräsentative und meist internationale Untersuchungen haben hingegen die Naturwissenschaften als globales Konstrukt im Fokus (Beaton et al., 1996; OECD, 2000; O'Sullivan et al., 1998). Eine Ausnahme bilden die TIMSS-Erhebungen in der Sekundarstufe II, in denen Kompetenzen in Physik getestet werden (Mullis, Martin, Robitaille & Foy, 2008). Da dieses

international weit verbreitete naturwissenschaftenübergreifende Denken in Deutschland strukturell kaum vorhanden ist, steht die Frage nach der Existenz disziplinspezifischer Kompetenzen versus einer übergreifenden naturwissenschaftlichen Kompetenz im Raum. Die Prüfung dieser aus deutscher Sicht interessanten Fragestellung kann in diesem Rahmen nur auf eine der naturwissenschaftlichen Disziplinen – die Biologie – beschränkt werden. Dieses Vorgehen steht im Einklang mit den Herangehensweisen anderer Projekte im Bereich Kompetenzmodellierung. Auch im DFG-Schwerpunktprogramm Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen werden in Einzelprojekten detailliert Kompetenzfacetten unterschiedlicher Disziplinen modelliert (Klieme & Leutner, 2006) – z. B. in den Naturwissenschaften die „Bewertungskompetenz für systematisches Entscheiden in komplexen Situationen Nachhaltiger Entwicklung“ (Bögeholz/Hasselhorn/Watermann).

Da sie einen wesentlichen Beitrag zum individuellen Selbst- und Weltverständnis von Personen leistet (Gräber & Nentwig, 2002; KMK, 2005a), wird die Disziplin Biologie zur näheren Untersuchung herangezogen. In der Biologie – der Lehre vom Leben – findet die Auseinandersetzung mit dem Lebendigen statt, und die Natur wird interpretiert (Treviranus, 1802). Zudem ist ein Denken in unterschiedlichen Systemen (z. B. System Erde, System Mensch, System See) notwendig. Der Mensch wird als Teil und Gegenüber der Natur betrachtet. Damit einhergehend werden umweltverträgliches, gesundheitsbewusstes und nachhaltiges Handeln gefördert (Gräber & Nentwig, 2002). Dem Individuum wird es ermöglicht, die es umgebende natürliche, kulturelle und technische Welt zu verstehen und zu erklären – aber auch die Risiken und Chancen vergangener, momentaner und zukünftiger Entwicklungen für das eigene Leben und die Gesellschaft zu erkennen (DeBoer, 2000; KMK, 2005a; Winick et al., 2008).

Wie bereits angedeutet, ist die Literaturlage zur Definition von Kompetenz in Biologie überschaubar und nicht ausreichend. Daher werden die bildungspolitisch relevanten Instrumente Bildungsstandards und Rahmenlehrpläne herangezogen. Mittels deren Verortung wird deutlich, wie Bildung im Fach Biologie in Deutschland zum momentanen Zeitpunkt zu gestalten ist und wie sie in Zukunft vorangetrieben werden soll. Die beiden Instrumente werden aufgrund ihrer Umfänglichkeit nur skizziert und als weitere Basis für die Definition von Kompetenz in Biologie einbezogen.

1.3.1 Kompetenz in Biologie – Fakten- und Handlungswissen in der schulischen Praxis

Für die Definition einer disziplinbezogenen Kompetenz ist es unabdingbar, schulbezogene Auffassungen dieser Kompetenz einzubinden. Zunächst werden daher die zunehmend an Bedeutung gewinnenden bundesweit gültigen Bildungsstandards für die Sekundarstufe I im Fach Biologie vorgestellt. Diese sind weder flächendeckend noch in der Tiefe im deutschen Fachunterricht

implementiert (Pant, Vock, Pöhlmann & Köller, 2008). Für die weiterführende Sekundarstufe II hat die KMK 1989 Einheitliche Prüfungsanforderungen in der Abiturprüfung (EPA) Biologie erlassen, welche 2004 in einer neuen Fassung vorgelegt wurden (KMK, 2004). Sie spiegeln die Struktur der im Folgenden vorgestellten Bildungsstandards für die Sekundarstufe I wider und entwickeln sie weiter. Dieser Aufbau soll eine kontinuierliche Kompetenzentwicklung in Biologie gewährleisten. Daran anschließend folgt die Vorstellung länderspezifischer Schulcurricula für das Fach Biologie am Beispiel der Bundesländer Berlin und Bayern. Sie stellen das unmittelbar mit Unterricht in Verbindung stehende politische Instrument dar. Die Vorstellung wird ausführlich vorgenommen, da sie nicht nur einen Einblick in die Gewichtung von Konzept- und Prozessaspekten in bildungstheoretischen und -praktischen Zusammenhängen aufzeigt, sondern zudem weitere theoretische Unterfütterungen für die anschließende Definition von Kompetenz in Biologie liefert.

1.3.1.1 Bildungsstandards im Fach Biologie

Zur Bündelung der aus internationalen *large-scale assessment* Studien entstandenen Entwicklungen und Forderungen wurde 2006 die Gesamtstrategie der KMK zum Bildungsmonitoring publiziert (KMK, 2006). Als eine dieser Entwicklungen kann die bundesweite Einführung der Bildungsstandards – 2004 für den Mittleren Schulabschluss in den Fächern Biologie, Physik und Chemie – gesehen werden (KMK 2005a, 2005b, 2005c). Die KMK hat mit diesem Schritt die Naturwissenschaften als ein die Gesellschaft prägendes Fachgebiet identifiziert.

In den Dokumenten zu den Bildungsstandards werden Erwartungen an den Entwicklungsstand von Kindern und Jugendlichen zu einem bestimmten Zeitpunkt innerhalb ihrer Bildungskarriere formuliert. Bezogen auf die naturwissenschaftliche Kompetenz konkretisieren die Bildungsstandards die Hauptziele in Form von Kompetenzanforderungen (Klieme & Leutner, 2006). Im deutschen Kontext besteht die Schwierigkeit darin, ein einheitliches Konzept der naturwissenschaftlichen Kompetenz bei gleichzeitiger Berücksichtigung der spezifischen Beiträge der Fächer Biologie, Physik und Chemie zu entwickeln. Die Bundesländer wurden angehalten, die Standards zu Beginn des Schuljahres 2005/06 im Unterricht zu implementieren.

Die gemeinsame Grundlage der Bildungsstandards für die naturwissenschaftlichen Disziplinen sind die vier übergreifenden Kompetenzbereiche Fachwissen, Erkenntnisgewinnung, Bewertung und Kommunikation. Für das Fach Biologie werden die spezifischen Elemente der Kompetenzbereiche Fachwissen und Erkenntnisgewinnung näher erläutert, da sie grundlegend für die untersuchten Teilbereiche von Kompetenz in Biologie sind und analog zur Unterscheidung von Konzept- und Prozesswissen der Konzeptualisierungen zu *scientific literacy* gesehen werden können. Tabelle 5 gibt einen ersten Überblick über die Inhalte der vier Kompetenzbereiche.

Tabelle 5: Inhalte der vier Kompetenzbereiche der Bildungsstandards im Fach Biologie

Kompetenzbereich	Inhalt
Fachwissen	Lebewesen, biologische Phänomene, Begriffe, Prinzipien, Fakten kennen und den Basiskonzepten zuordnen
Erkenntnisgewinnung	Beobachten, Vergleichen, Experimentieren, Modelle nutzen und Arbeitstechniken anwenden
Kommunikation	Informationen sach- und fachbezogen erschließen und austauschen
Bewertung	Biologische Sachverhalte in verschiedenen Kontexten erkennen und bewerten

Aus „Bildungsstandards Biologie“, von KMK, (2005a), S. 7.

Der *Kompetenzbereich Fachwissen* bildet die Fachinhalte und somit das grundlegende zu erarbeitende Fakten- und Konzeptwissen des Faches Biologie ab. In allen drei akademischen Disziplinen, so auch in der Biologie, wird diese Dimension durch Basiskonzepte repräsentiert. Die Basiskonzepte für das Fach Biologie sind (a) System, (b) Struktur und Funktion sowie (c) Entwicklung. Die inhaltliche Ausgestaltung kann Tabelle 30 im Anhang entnommen werden.

Die Bildungsstandards orientieren sich also an einer geringen Anzahl von basalen fachlichen Konzepten, die übergreifende Ideen vereinen. So beinhaltet das Basiskonzept System in der Biologie Biosysteme. Dazu gehören wiederum globale Systeme wie der Mensch oder das Ökosystem sowie Teilsysteme wie die Zelle. Diese Konzepte werden auch durch die formulierten Regelstandards für den Kompetenzbereich Fachwissen verdeutlicht. In der folgenden Tabelle 6 werden zwei Standards pro Basiskonzept exemplarisch aufgezeigt.

Tabelle 6: Exemplarische inhaltliche Ausgestaltung der Standards des Kompetenzbereichs Fachwissen nach Basiskonzepten

Die Schülerinnen und Schüler ...
F1 System
F1.1 ... verstehen die Zelle als System,
F1.6 ... stellen einen Stoffkreislauf sowie den Energiekreislauf in einem Ökosystem dar,
F2 Struktur und Funktion
F2.1 ... beschreiben Zellen als strukturelle und funktionelle Grundbaueinheiten von Lebewesen,
F2.5 ... beschreiben die strukturelle und funktionelle Organisation im Ökosystem,
F3 Entwicklung
F3.1 ... erläutern die Bedeutung der Zellteilung für Wachstum, Fortpflanzung und Vermehrung,
F3.4 ... beschreiben ein Ökosystem in zeitlicher Veränderung.

Aus „Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss“, von KMK, 2005a, S. 13 f..
Anmerkungen. Vollständige Darstellung im Anhang Tabelle 32.

Bezogen auf das Beispiel Zelle und Ökosystem sollen die Schülerinnen und Schüler die Thematiken sowohl verstehen und erklären als auch deren Wechselwirkungen mit anderen Systemen und ihrer Umwelt beschreiben. Eine dritte Ebene stellen die Anforderungsbereiche für jeden Kompetenzbereich dar. Diese sind für das Fachwissen Wiedergabe, Anwendung und Transfer.

Der *Kompetenzbereich Erkenntnisgewinnung* bildet eine der Handlungsdimensionen von Kompetenz in Biologie ab. Diese umfasst die grundlegenden wissenschaftsmethodischen Verfahren kriteriengeleitete Beobachtung biologischer Phänomene, hypothesengeleitetes Experimentieren, kriterienbezogenes Vergleichen und die Modellbildung. Auch deren inhaltliche Ausgestaltung wird im Anhang dargestellt. Schülerinnen und Schüler sollen Fragestellungen formulieren und Hypothesen aufstellen können, sowie eine Untersuchung planen und durchführen, die Daten analysieren und interpretieren können. Sie sollen also naturwissenschaftliche Untersuchungsmethoden anwenden können. Die entsprechenden Formulierungen für den Kompetenzbereich Erkenntnisgewinnung werden in Tabelle 7 exemplarisch aufgeführt.

Tabelle 7: Exemplarische inhaltliche Ausgestaltung der Standards des Kompetenzbereichs Erkenntnisgewinnung

Die Schülerinnen und Schüler ...
E1 mikroskopieren Zellen und stellen sie in einer Zeichnung dar,
E3 analysieren die stammesgeschichtliche Verwandtschaft bzw. ökologisch bedingte Ähnlichkeit bei Organismen durch kriteriengeleitetes Vergleichen,
E6 planen einfache Experimente, führen die Experimente durch und/oder werten sie aus,
E9 wenden Modelle zur Veranschaulichung von Struktur und Funktion an.

Aus „Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss“, von KMK, 2005a, S. 14.
Anmerkungen. Vollständige Darstellung im Anhang Tabelle 33.

In den Bildungsstandards werden also sowohl Aspekte des Fachwissens und Prozessaspekte von Kompetenz in Biologie berücksichtigt als auch der Anwendungsbezug hervorgehoben. Den kompetenzorientierten impulsgebenden Bildungsstandards stehen die praktisch angewendeten Schulcurricula der Bundesländer gegenüber. Die Implementation der Standards wird in der Mehrzahl der Bundesländer durch eine Überarbeitung der Curricula vorgenommen.

1.3.1.2 Schulcurricula am Beispiel der Bundesländer Berlin und Bayern

Um einen Eindruck von aktueller Wissensvermittlung in deutschen Schulen zu gewährleisten, werden im Folgenden die Schulcurricula vorgestellt. Schulcurricula stellen intendierte Curricula dar. Diese unterscheiden sich von den implementierten und tatsächlich von den Schülerinnen und Schülern erhaltenen, den potentiellen und den erreichten Curricula (Köller, Baumert & Bos, 2001; Mullis et al., 2003). Sie stehen für Inhalte, die die bildungspolitischen Entscheidungsträger für deutsche Schülerinnen und Schüler anstreben. Hierbei soll ein besonderes Augenmerk auf die Einbeziehung von Aspekten gelenkt werden, die über die Vermittlung von Fachwissen hinausgehen.

Im föderal organisierten Deutschland werden den einzelnen Bundesländern eigene Curricula entwickelt. Da eine Aufbereitung dieser Fülle an Material den Rahmen dieser Arbeit sprengen würde und dem Zweck des exemplarischen Aufzeigens von Bildungsinhalten zur Definition von Kompetenz in Biologie nicht angemessen wäre, sollen an dieser Stelle beispielhaft die Lehrpläne von zwei Bundesländern aufgeführt werden: Berlin und Bayern. Beide Bundesländer sind in der Stichprobe der dieser Arbeit zu Grunde liegenden Untersuchung enthalten (s. Seite 101).

Schulcurriculum Berlin

Der Berliner Rahmenlehrplan für das Fach Biologie für die Jahrgangsstufe 7 bis 10 wurde 2006 von der SBJs veröffentlicht (SBJs, 2006). Schon ein Blick auf das Inhaltsverzeichnis zeigt, dass der Rahmenlehrplan im Sinne der Bildungsstandards aufgebaut ist; neben der Fachkompetenz wurden

die drei weiteren Kompetenzteilbereiche Erkenntnisgewinnung, Bewertung und Kommunikation als durch den Unterricht zu entwickelnde Teilkompetenzen aufgenommen. Auf die für die vorliegende Arbeit entscheidenden Kompetenzteilbereiche Erkenntnisgewinnung und Fachwissen soll im Folgenden eingegangen werden.

Das Fachwissen wird durch die bereits bekannten Basiskonzepte aus den Bildungsstandards widerspiegelt. Aus der Vielfalt der möglichen Themen können Inhalte im Unterricht nur exemplarisch behandelt werden. In der folgenden Tabelle 8 werden die zu behandelnden Themen für die mittleren Standards aufgelistet, welche für die Realschule und die E-Kurse der Gesamtschule gelten.

Tabelle 8: Mittlere Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für das Basiskonzept Entwicklung

Die Schülerinnen und Schüler ...
... beschreiben die menschliche Entwicklung von der Zygote bis zum Tod.
... erläutern die Individualentwicklung von Organismen im Überblick.
... unterscheiden verschiedene Formen der Fortpflanzung.
... analysieren und erklären stammesgeschichtliche Verwandtschaft von Organismen.
... erklären Organsysteme und deren Anpasstheit im Zusammenhang zur Evolutionstheorie.
... untersuchen und erklären den Verlauf stammesgeschichtlicher Entwicklung an ausgewählten Lebewesen.
... erklären die Variabilität von Lebewesen.
<i>Anmerkungen.</i> Mittlere Standards im Berliner Rahmenlehrplan wurden für die Realschule und die E-Kurse der Gesamtschule, die einfachen Standards für die Hauptschule und die G/A-Kurse der Gesamtschule und die erweiterten Standards für das Gymnasium und die F-Kurse der Gesamtschule entwickelt. Die vollständige Tabelle 34 für die mittleren Standards in allen Basiskonzepten sowie die Tabelle 35 für alle Niveaustufungen für das Basiskonzept Entwicklung befinden sich im Anhang. Aus „Rahmenlehrplan für die Sekundarstufe I“, von SBJs, 2006, S. 18.

Der Bereich Erkenntnisgewinnung des Berliner Schulcurriculums stellt eine der Handlungsdimensionen dar. Diese umfasst experimentelles und theoretisches Arbeiten. Hierzu gehören die Beobachtung, der Vergleich, das Experimentieren sowie die Modellbildung. Die Schülerinnen und Schüler sollen diese Fertigkeiten in einem problemorientierten Unterricht erlangen und können nur gemeinsam mit den weiteren Dimensionen erworben werden (SBJs, 2006). In der Tabelle 9 werden die mittleren Standards zum Kompetenzbereich Erkenntnisgewinnung aufgelistet.

Tabelle 9: Mittlere Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für den Kompetenzbereich Erkenntnisgewinnung

Die Schülerinnen und Schüler ...

- ... ermitteln Verwandtschaft bzw. ökologisch bedingte Ähnlichkeit bei Organismen durch kriteriengeleitetes Vergleichen.
 - ... erklären die Speicherung und Weitergabe genetischer Information auch unter Anwendung geeigneter Modelle.
 - ... führen Experimente zu Fotosynthese und Atmung durch und werten sie aus.
 - ... nutzen ggf. den Computer zur Darstellung und Auswertung von Messreihen oder zur Simulation biologischer Abläufe.
 - ... bestimmen den eigenen Lerntyp, entwickeln und erläutern individuelle Lernstrategien.
-

Aus „Rahmenlehrplan für die Sekundarstufe I“, von SBJS, 2006, S. 19

Der im Curriculum folgende Abschnitt der Themen und Inhalte des Biologieunterrichts teilt diesen in verschiedene inhaltliche Dimensionen ein und fokussiert auf (a) einheimische Lebewesen und Wechselwirkungen, (b) Bau und Leistungen des menschlichen Körpers I+II, (c) Liebe, Partnerschaft und Sexualität, (d) einheimische Pflanzen und ihre Bedeutung im Stoffkreislauf, (e) Grundlagen der Vererbung und (f) Abstammungslehre. Zur Verdeutlichung wird in Tabelle 10 der erste Themenbereich der Doppeljahrgangsstufe 9/10 in seiner Ausgestaltung – erneut exemplarisch anhand der mittleren Standards – zitiert.

Tabelle 10: Kompetenzbezüge der Inhalte von Biologieunterricht im Berliner Schulcurriculum am Beispiel des Themenfeldes *Einheimische Pflanzen und ihre Bedeutung im Stoffkreislauf*

Einheimische Pflanzen und ihre Bedeutung im Stoffkreislauf

Die Schülerinnen und Schüler ...

- ... vergleichen die Grundbaupläne der Blüten einheimischer Pflanzenfamilien.
 - ... erklären die Anpasstheit der Pflanzenorgane an unterschiedliche Umweltbedingungen.
 - ... begründen die Bedeutung der Pflanze am Beispiel des Kohlenstoffkreislaufs.
-

Aus „Rahmenlehrplan für die Sekundarstufe I“, von SBJS, 2006, S. 36.

Die Tätigkeit Vergleichen stellt bereits eine erkenntnistheoretische Handlung der Schülerinnen und Schüler dar. Das Erklären und Begründen ist im Konzeptwissen anzusiedeln. Mit diesem Beispiel konnte ein überarbeitetes Curriculum aufgezeigt werden, welches bereits Prozessaspekte von Kompetenz in Biologie beinhaltet (z. B. die Schülerinnen und Schüler führen Experimente durch). Das nun folgende Beispiel aus dem Bundesland Bayern folgt eher den traditionellen Ansichten.

Kapitel 1

Schulcurriculum Bayern

Der Fachlehrplan für Biologie für das Gymnasium und die Realschule wurde in Bayern 2004 und 2003 veröffentlicht. In Tabelle 11 werden die Unterrichtsinhalte der Jahrgangsstufe 10 im Gymnasium in verkürzter Form dargestellt.

Kapitel 1

Tabelle 11: Unterrichtsinhalte im Fach Biologie der Jahrgangsstufe 10 im Bayerischen Curriculum

Themenbereich	Inhalt
Realschule	
Üben und vertiefen, Schülerübungen	
Genetik	Weitergabe von Erbinformationen Gene und Merkmale Mendelsche Regeln Angewandte Genetik
Verantwortliche Elternschaft	Eltern werden Die ersten Lebensjahre
Biologie und kulturelle Evolution des Menschen	Stammesgeschichte des Menschen Kulturelle Evolution
Gymnasium	
Stoffwechsel des Menschen	Ernährung und Verdauung Atmung und Blutkreislauf Stoffwechsel in der Zelle
Bau, Funktionsweise und Schädigung von inneren Organen	
Grundlegende Wechselbeziehungen zwischen Lebewesen	Die Umwelt eines Lebewesens Beziehungen zwischen Lebewesen Aufbau und Merkmale eines Ökosystems der gemäßigten Breiten an einem konkreten Beispiel Bedeutung und Gefährdung von Ökosystemen
Angewandte Biologie	Biotechnologie Landwirtschaft Medizin

Aus „Jahrgangsstufenlehrplan für Biologie an Gymnasien in Bayerischen Schulen“, von Staatsinstitut für Schulqualität und Bildungsforschung München [ISB], 2004 <http://www.isb-gym8-lehrplan.de/content/serv/3.1.neu/g8.de/index.php?StoryID=26227> und

Aus „Jahrgangsstufenplan für Biologie an Realschulen in Bayerischen Schulen, von ISB, 2003, <http://www.isb.bayern.de/isb/download.aspx?DownloadFileID=b4e7bb42889dba9a7659f0c38ed9262b>.

Anmerkungen. Üben und vertiefen, Schülerübungen ist ein Platzhalter von ca. 12 Stunden, der den Lehrerinnen und Lehrern die Möglichkeit gibt (a) Themenbereiche zu vertiefen, (b) Schülerübungen durchzuführen oder (c) aktuelle Schülerinteressen und Themen aus dem Tagesgeschehen in den Unterricht einzubringen.

Es fällt auf, dass in diesem älteren Lehrplan weitaus weniger auf die Handlungsdimension eingegangen wird als im Lehrplan des Bundeslandes Berlin; die Kompetenzbereiche Bewertung und

Kommunikation fehlen vollständig. Zudem werden wesentlich kleinschrittiger im Unterricht zu behandelnde fachliche Themen aufgezeigt, sodass schwer vorstellbar ist, dass Lehrkräfte den Handlungsaspekt im Unterricht hervorheben können.

Die Darstellung der Bildungsstandards für das Fach Biologie und der Schulcurricula der Bundesländer Berlin und Bayern hat gezeigt, dass Kompetenz in Biologie auch in der praktischen Anwendung Fachwissen- und anwendungsbezogene Aspekte beinhalten kann. In den Naturwissenschaften und in der Biologie erhält in diesem Zusammenhang die Erkenntnisgewinnung eine herausgehobene Stellung. Sie umfasst, wie an den Schulcurricula deutlich gemacht wurde, in der schulischen Praxis meist naturwissenschaftliche Arbeitsweisen. Die vorangegangenen theoretischen Ausführungen zeigen jedoch, dass auch die Denkweisen, die das naturwissenschaftliche Forschen bestimmen, ein zu beachtender Aspekt von Kompetenz in Biologie sind (Roberts & Gott, 1999). Im Sinne einer zukunftsweisenden Definition wird Kompetenz in Biologie im folgenden Abschnitt im Lichte der vorangegangenen Betrachtung des Wissenstandes definiert.

1.3.1.3 Von *scientific literacy* zu einer Definition von Kompetenz in Biologie

Wie aus den berichteten Konzeptualisierungen von *scientific literacy* sowie aus den exemplarischen Schulcurricula und den Bildungsstandards für das Fach Biologie ersichtlich wurde, sollte eine Definition von Kompetenz in Biologie nicht nur das notwendige Fachwissen beinhalten. Wenn sie funktional, also anwendungsorientiert ausgerichtet sein soll, muss naturwissenschaftliche Kompetenz bzw. Kompetenz in Biologie vielmehr aus mehreren Teilkomponenten zusammengesetzt sein. Dies bedeutet zum einen, dass *scientific literacy* und *biological literacy* einige Schülerinnen und Schüler darauf vorbereiten sollte, aktive Biologinnen bzw. Biologen zu werden. Zum anderen muss das Ziel der Vermittlung von Kompetenz in Biologie sein, alle Schülerinnen und Schüler mit den notwendigen Kompetenzen auszustatten, welche sie dazu befähigen, eigenständige Entscheidungen zu alltagsrelevanten naturwissenschaftlichen und biologischen Thematiken, wie z. B. Umweltverschmutzung und -schutz oder Impfschutz (Osborne, 2007) zu treffen. Kompetenz in Biologie, die ausschließlich die konzeptuellen Aspekte beinhaltet, wird dieser lebensnahen Ausrichtung nicht gerecht. Bezogen auf generelle Wissensbestände liegen einschlägige Theorien zur Unterteilung von Wissen in eine Prozess- und eine Wissenskomponente vor, welche noch aufgegriffen werden (s. Seite 50). Diese werden in empirischen Untersuchungen insbesondere im Fach Mathematik und vereinzelt auch in den Naturwissenschaften auf die disziplinspezifischen Kompetenzen übertragen. In der vorliegenden Arbeit werden daher diese beiden Komponenten untersucht und finden in der folgenden Definition von Kompetenz in Biologie ihren Niederschlag.

Definition von Kompetenz in Biologie

Kompetenz in Biologie beinhaltet bei Individuen verfügbare bzw. erlernbare kognitive, kontextspezifische Leistungsdispositionen, die sie dazu befähigen, biologiebezogene Probleme in variablen Situationen zu lösen. Angelehnt an den Begriff der *scientific literacy* setzt sie sich aus einer Konzept- und einer Prozesskomponente zusammen.

Die Konzeptkomponente beinhaltet hierbei das fachliche Wissen – inhaltlich organisiert über die Konzepte (a) System, (b) Struktur und Funktion sowie (c) Entwicklung – und das Verständnis über die Zusammenhänge zwischen fachlichen Fakten. Die Prozesskomponente spiegelt den Anwendungsbezug wider. Sie umfasst das Verständnis über die Biologie und die erkenntnistheoretischen Fähigkeiten im Zusammenhang mit der Wissensgenerierung in dieser Disziplin. Sie beinhaltet die grundlegenden wissenschaftsmethodischen Herangehensweisen naturwissenschaftliche Untersuchungen, den Umgang mit Modellen und den reflektierten Umgang mit der Biologie und ihren Methoden.

Diese Definition kann und soll keine umfassende Definition von Kompetenz in Biologie darstellen, weitere Teilbereiche wie das Wissen um die adressatengerechte Kommunikation oder die auf den Naturwissenschaften basierende Meinungsbildung wären in einer Definition dieser Kompetenz vorstellbar. Da in der vorliegenden Arbeit die Teilbereiche Konzept- und Prozesswissen untersucht werden sollen, wird auch die Definition von Kompetenz in Biologie daraufhin fokussiert. In der Definition wird die Auffächerung von Kompetenz in inhaltliche und konzeptuelle Aspekte einerseits und in Prozesse und Prozeduren andererseits herausgearbeitet. Beispielhaft für konzeptuelle Aspekte ist das Wissen um die Rolle von Glukose bei der Photosynthese oder von Sauerstoff bei der Atmung; für Prozessaspekte sind der Aufbau und das Design einer Untersuchung oder die Interpretation von Daten beispielhaft. Sie steht im Einklang mit Konzeptualisierungen internationaler *large-scale assessment* Studien, den Bildungsstandards für die naturwissenschaftlichen Fächer und darauf aufbauenden Schulcurricula. Bei der Beschreibung disziplinspezifischer Kompetenzen anhand von Kompetenzmodellen spielen sowohl inhaltliche Aspekte als auch Elemente der kognitiven Entwicklung eine Rolle. So sollten diese Kompetenzmodelle strukturelle Ähnlichkeiten mit kognitionspsychologischen Theorien aufweisen (Bernholt et al., 2009). Die aufgezählten Konzeptualisierungen mit der grundlegenden Unterscheidung einer Wissens- von einer Prozesskomponente sind beeinflusst von kognitionspsychologischen Unterscheidungen von Wissensformen. Diese werden im folgenden Kapitel aufgezeigt, um der erarbeiteten Dimensionalisierung von Kompetenz in Biologie ein weiteres theoretisches Fundament zu geben.

Kapitel 1

Bereits die erarbeitete Definition von Kompetenz in Biologie enthält theoretisch begründete Vermutungen zur internen Struktur des Konstrukts. Sie spricht damit einen von sechs Aspekten der Konstruktvalidität an – den strukturellen Aspekt (Messick, 1995). Daher werden im Folgenden der Begriff Validität und die unterschiedlichen Aspekte näher erläutert. Die Unterscheidung von Kompetenz in Biologie in die Teilbereiche Konzept- und Prozesswissen wird darin verortet und die Aufnahme der weiteren Validitätsaspekte in die vorliegende Arbeit strukturiert.

2 Grundlagen zur Validierung eines Tests zu Kompetenz in Biologie

Der Begriff Validität beschreibt eine integrierende Bewertung. Bewertet wird die Angemessenheit und Güte von tatsächlichen und potenziellen Interpretationen und Maßnahmen auf Basis von Testwerten und anderen Erhebungsmitteln. Mit Hilfe von empirischen Belegen und theoretischen Annahmen kann diese Angemessenheit und Güte bestimmt werden (Hartig et al., 2007; Messick, 1995). Im Zuge der Validierung eines Tests wird also einerseits das Ausmaß gemessen, mit dem der Test das Konstrukt misst, welches er messen soll. Andererseits muss eine Validierung auch die Konsequenzen der unterschiedlichen intendierten und nicht intendierten Interpretationen des Tests –also der Nutzung – einbeziehen (Messick, 1995).

Die Notwendigkeit der Validierung und die Definition des Begriffs Validität ergab sich historisch nach den ersten Entwicklungen von empirischen Tests zu Anfang des letzten Jahrhunderts. Die Validierung sollte die akurate Verwendung dieser Tests gewährleisten und rechtfertigen (Zumbo, 2009). Seither hat die Definition von Validität einen starken Wandel durchlaufen. Zum einen rührt dieser Wandel von der Entwicklung und Anwendung neuer Testverfahren, zum anderen führten verschiedene theoretische Überlegungen zu deren Erweiterung.

In den ersten Definitionen von Validität stand die Kriteriumsvalidität, also die Bestimmung des Zusammenhangs „zwischen den individuellen Ausprägungen des Kriteriums und den entsprechenden individuellen Testwerten“ (Hartig, Frey & Jude, 2007, S. 137), im Vordergrund. Durch neue statistische Verfahren entstand die Idee der Konstruktvalidität. Nun spielte auch die interne Struktur eines Konstrukts – und damit Aspekte der Inhaltsvalidität und die inkrementelle Validität – in die Bewertung der Validität eines Tests eine Rolle. (Lissitz, 2009). Im Zuge der inkrementellen Validierung können neue Testteile auf ihre Passung zum bereits bestehenden Test bewertet werden (Lissitz, 2009). Diese Aufsplittung führte im weiteren Verlauf zur Ausdifferenzierung aber auch Zerfaserung des Validitätskonzepts. Die unterschiedlichen Validitätsformen wurden in die übergreifenden Validitätstypen Inhalts-, Kriteriums- und Konstruktvalidität unterteilt (Cronbach & Meehl, 1955) und durch Messicks vereintes Validitätskonzept zusammengeführt (Baartman et al., 2007; Messick, 1995). Mit dieser neuen Auffassung erfährt der Begriff Validität weiterhin eine Ausdehnung zu früheren Konzeptualisierungen. Mit der Ausdifferenzierung und Vereinigung der Definition von Validität ging immer auch eine theoretische Auseinandersetzung mit dem Begriff einher. Diese führte zu der Erweiterung des Konzepts um die Bewertung der intendierten und nicht-intendierten Konsequenzen eines Tests.

Nach heutiger Auffassung umfasst Konstrukt- bzw. Testvalidität „die empirischen Befunde und Argumente, mit denen die Zuverlässigkeit der Interpretation von Testergebnissen im Sinne erklärender Konzepte, die sowohl die Testergebnisse selbst als auch die Zusammenhänge der Testwerte mit anderen Variablen erklären, gestützt wird“ (Hartig et al., 2007, S. 145). Mit dieser Definition wird die traditionelle Auffassung von Validität erweitert und die übergreifende Konstrukt- bzw. Testvalidität in die sechs Aspekte (a) inhaltliche Relevanz und Repräsentativität, (b) substanzielle Theorien, Prozessmodelle und Aufgabenbearbeitungsprozesse, (c) Angemessenheit der Bewertungsverfahren, (d) Verallgemeinerung und Grenzen der Interpretation und Verwendung eines Tests, (e) Zusammenhang mit externen Variablen sowie (f) Konsequenzen der Interpretation und Verwendung differenziert (Messick, 1995). Der Inhaltsaspekt nimmt Bezug auf Evidenzen zur Relevanz des Testinhalts, zur Repräsentativität und zur technischen Qualität des Tests. Mit dem substanziellen Aspekt werden theoretische Erklärungen für beobachtete Konsistenzen in der Beantwortung des Tests gefordert. Die Testung soll damit ermöglichen, das Verhalten und die Wissensanwendung in der Realität zu reflektieren (Baartman et al., 2007). Der Strukturaspekt betrifft die Stimmigkeit der Struktur der Testwerte mit der Struktur des getesteten Konstrukts. Der Aspekt der Generalisierbarkeit umschreibt das Maß, in dem Eigenschaften der Testwerte und -interpretationen auf Populationsgruppen, Settings und Aufgaben übertragen werden können. Der Externalitätsaspekt beinhaltet konvergente und diskriminante Evidenzen im Sinne der *multitrait-multimethod* Vergleiche (Campbell & Fiske, 1959). Hierunter fällt auch das vieldiskutierte Aufspannen eines nomologischen Netzes um das zu erfassende Konstrukt (Borsboom, Mellenbergh & van Herden, 2004; Cronbach & Meehl 1955; Hartig et al., 2007). Ein nomologisches Netzwerk besteht aus Axiomen und Korrespondenzregeln. Aus diesen entsteht ein System von Gesetzen, welches die hinter einem Test liegende Theorie konstituiert (Cronbach & Meehl, 1995). Dieses nomologische Netz ist untersuchbar und somit die Validität eines Tests einschätzbar. Die sehr technischen Begrifflichkeiten zur Beschreibung dieses Netzes dürfen nicht darüber hinwegtäuschen, dass deren Ergebnisse nur Hinweise für die Validität eines Tests geben können (Cronbach & Meehl, 1995). Der Aspekt der Konsequenzen schließlich bezieht sich auf Wertimplikationen von Testwertinterpretationen als Basis für Maßnahmen sowie auf tatsächliche und potenzielle Konsequenzen. Diese sechs Aspekte der Validität können entweder der Klärung der Verallgemeinerbarkeit eines Testergebnisses, dessen kausaler Erklärung, dessen Extrapolierung auf andere Bereiche oder den Entscheidungen aufgrund von Testergebnissen dienen.

Die *American Educational Research Association* (AERA), *American Psychological Association* (APA) und der *National Council on Measurement in Education* (NCME) haben die vier Quellen der Validität Inhalt, Antwortprozesse, interne Struktur und Konsequenzen des Testens postuliert (AERA, 2004), in

welche die sechs Aspekte integriert werden können. In Tabelle 12 wird diese Integration vorgenommen und die vorliegende Arbeit in diesem Rahmen verortet.

Tabelle 12: Integration der Validitätsaspekte von Messick (1995) und der Quellen von Validität der AERA, APA & NCME (2004) sowie Einordnung der vorliegenden Arbeit

Aspekte nach Messick (1995)	Quelle nach AERA, APA & NCME (2004)	Einordnung der vorliegenden Arbeit
Inhaltliche Relevanz und Repräsentativität	Inhalt, interne Struktur	X
Substanzielle Theorien, Prozessmodelle und Aufgabenbearbeitungsprozesse	Inhalt / Antwortprozesse	
Angemessenheit der Bewertungsverfahren	Antwortprozesse	
Verallgemeinerung und Grenzen der Interpretation und Verwendung	Konsequenzen des Tests	(X)
Zusammenhang mit externen Variablen	Inhalt	X
Konsequenzen der Interpretation und Verwendung	Konsequenzen des Tests	

Anmerkungen. AERA = American Educational Research Association, APA = American Psychological Association, NCME = National Council on Measurement in Education (NCME).

Die Erweiterung des Begriffs Validität führt zu einem Kontinuum, welches durch die Skizzierung der beiden Pole dieses Kontinuums beschrieben werden kann. Der eine Pol ist charakterisiert durch das Vorantreiben der Erweiterung des Konzepts Validität um die Interpretationen und die Anwendungen eines Tests (z. B. Kane, 2002). Der andere Pol des Kontinuums wird durch die unterschiedlichen Herangehensweisen an die Validierung eines Tests bestimmt. Diese Herangehensweisen fokussieren die inhaltliche Validierung einer zu messenden Theorie (Borsboom et al., 2004; Lissitz & Samuelson, 2007). Die Verwendung des Begriffs Kontinuum macht deutlich, dass eine Vereinigung beider Pole möglich ist. Die vorliegende Arbeit liegt in diesem Kontinuum eher auf Seiten der inhaltlichen Validierung. Es wird versucht, Konzept- und Prozesswissen und deren Testung einzuordnen. Die intendierten und nicht-intendierten Konsequenzen werden nur partiell tangiert.

Unter dem Begriff Validität werden also unterschiedliche Prüfverfahren und Kriterien zusammengefasst, die zur Verortung eines Tests bezogen auf die dahinter liegende Theorie und die Interpretationen eingesetzt werden können (Hartig et al., 2007). Wenn ein Test aus

unterschiedlichen Blickwinkeln betrachtet wird und die Resultate aus diesen Blickwinkeln positiv ausfallen, kann der Test (für die intendierte Nutzung) als valide angesehen werden (Lissitz, 2009). Aufgrund der Fülle an Verfahren, Kriterien und Perspektiven gilt es zunächst, die anstehende Validierung einzugrenzen, indem der Zweck eines Tests bestimmt wird. Der vorliegende neuentwickelte Test soll das Konstrukt Kompetenz in Biologie mit seinen Teilbereichen Konzept- und Prozesswissen abbilden. Zu den Teilbereichen bietet der vorangegangene theoretische Teil der Arbeit erste Anhaltspunkte zu deren korrelativen Zusammenhängen. Diese Hinweise werden im Folgenden mit Theorien der Kognitionspsychologie weiter unterfüttert. Diese Untersuchung der Teilbereiche bezieht sich auf das Erklären eines Testergebnisses aufgrund der inhaltlichen Struktur eines Konstrukts. Empirisch trennbare Dimensionen bestätigen dann die Annahmen über das getestete Konstrukt.

Um den vorliegenden Test für Kompetenz in Biologie umfassender zu validieren, müssen weitere der genannten Validitätsaspekte einbezogen werden. So sollte sichergestellt werden, dass der Test tatsächlich Kompetenz in Biologie abbildet und nur gering durch externe Konstrukte beeinflusst wird. Ein Test zu Kompetenz in Biologie sollte nicht zu stark durch Elemente wie z. B. Rechenfähigkeit oder sprachliche Kompetenzen beeinflusst sein. Eine Möglichkeit zur Validierung eines Tests ist das In-Beziehung-setzen des getesteten Konstrukts zu externen Kriterien. Von diesen Kriterien wird angenommen, dass sie entweder stark oder gering mit dem zu prüfenden Konstrukt korrelieren. Bei der konvergenten Validität wird davon ausgegangen, dass die Zusammenhänge hoch ausfallen. Im Gegensatz dazu steht die diskriminante Validität. Hier wird von möglichst geringen Korrelationen ausgegangen (Hartig et al., 2007). Im Sinne einer erweiterten Validitätsprüfung werden verschiedene disziplinübergreifende, psychologische und schulische Kriterien bezogen auf die drei naturwissenschaftlichen Disziplinen aber auch auf die Disziplinen Deutsch und Mathematik herangezogen. Die vergleichende Evaluation der Zusammenhänge gibt dann Aufschluss über die diskriminante und konvergente Validität des Kompetenztests in Biologie.

Diese Erweiterung führt zur Einführung von zwei neuen Blickwinkeln bezüglich der Bestimmung von Kompetenz in Biologie, die wichtige Beiträge zur Formulierung von Kompetenzmodellen leisten (Klieme & Leutner, 2006). Die Zusammenführung von Erkenntnissen unterschiedlicher Disziplinen wird insbesondere – aber nicht nur – in der Entwicklung von disziplinspezifischen Tests als sehr hilfreich angesehen (Fischer et al., 2005; Zimmerman, 2007). Zunächst wird die kognitionspsychologische Sichtweise auf Wissensstrukturen näher beleuchtet, um den Annahmen zur internen Struktur von Kompetenz in Biologie ein stärkeres Fundament zu geben. Anschließend wird die differentialpsychologische Sichtweise herangezogen, um die Hinzunahme von externen Kriterien zu begründen und einzuordnen.

2.1 Die Naturwissenschaften aus kognitionspsychologischer Perspektive

Aus kognitionspsychologischer Perspektive sind für die Verortung der Kompetenzstrukturen Wissenserwerbsprozesse und -strukturen von Relevanz. Unter Wissenserwerb wird „der Aufbau und die fortlaufende Modifikation von Wissensrepräsentationen“ (Steiner, 2001, S. 163) im Gedächtnis verstanden. Über Wissenserwerbsprozesse wird demzufolge Wissen in einer bestimmten Form im Gedächtnis abgespeichert. Beim Erwerb fachbezogenen Wissens (z. B. in den Naturwissenschaften) werden „elaborierte Symbolsysteme für Zwecke der Repräsentation wie auch der Kommunikation fachlicher Inhalte ... benutzt und deren Handhabung“ (Seel, 2003, S. 61) erlernt.

Eine Wissensrepräsentation ist ein Abbild der Wirklichkeit im Gedächtnis eines Individuums. Demnach werden im Gedächtnis Bedeutungen von der realen Welt in abstrakten kognitiven Strukturen gespeichert, es können aber auch Abstraktionen wie Handlungen und Werte repräsentiert werden. Nach dieser Auffassung von Wissen ist Wissenserwerb eine persönliche und individuelle Konstruktion des jeweiligen Sachverhaltes (Gerstenmaier & Mandl, 1995). In der Lehr-Lernforschung in den Naturwissenschaften wird für diese Ansicht zum Lernvorgang der moderate Konstruktivismus herangezogen (Duit, 1995; Leutner, 1997; Mintzes, Wandersee & Novak, 1997). Diese Form des Konstruktivismus, welche starke Bezüge zu den Kognitionswissenschaften aufweist, wurde im Rahmen der Debatte über den (moderaten) Konstruktivismus und der damit einhergehenden Auseinandersetzung mit dem *conceptual change* in der naturwissenschaftsdidaktischen Forschergemeinschaft geformt. Der Ursprung der konstruktivistischen Sichtweisen liegt im radikalen Konstruktivismus und ist beeinflusst von der Systemtheorie, dem sozialen Konstruktivismus und empirischen Befunden der Neurobiologie (Riemeier, 2007). Im moderaten Konstruktivismus wird Lernen als konstruktiv, selbstdeterminiert, individuell, sozial und situiert gesehen (ebd., 2007). Es wurden wichtige Aspekte der sozialen Eingebundenheit von Wissenskonstruktionen und die konstruktivistische Sichtweise integriert. Diese Integration führte zu dem *conceptual change* Modell von Posner, Strike, Hewson und Gertzog (1982). Mit diesem Modell wird über Konstruktionsprozesse, wie z. B. dem kognitiven Konflikt, versucht, von den vorhandenen Vorstellungen der Schülerinnen und Schüler zu fachlichen Vorstellungen überzugehen (Duit, 1995). In der Auseinandersetzung mit dem *conceptual change* von Schülerinnen und Schülern werden das Wissen und die Vorstellungen berücksichtigt, die die Schülerinnen und Schüler in den Unterricht mitbringen (Krüger, 2007). Bei dem konstruktivistischen Blick auf Lehren und Lernen werden Schülerinnen und Schüler, die Lehrkraft, das Wissen sowie der Lernprozess als den Wissenserwerb konstruierende Koordinaten gesehen (Posner et al., 1982).

Unter den Kognitions- und Entwicklungspsychologinnen und -psychologen sowie in der Pädagogischen Psychologie divergieren nun die Meinungen und Theorien darüber, auf welche Art

diese kognitiven Strukturen im Gedächtnis abgespeichert werden, wie also die Struktur von Wissen zu fassen ist (z. B. Anderson, 1983; Baddeley, 1992). Diese divergierenden Meinungen führen zu unterschiedlichen Ordnungssystemen, sodass auch heute noch verschiedene Auffassungen zur Struktur von Wissensarten verbreitet sind. Im Laufe der Wissensforschung wurden unzählige Arten und Strukturierungen von Wissen angenommen. Konsens besteht darüber, dass das Gedächtnis aus einzelnen Subsystemen zusammengesetzt ist. Folgend werden exemplarisch einige Ordnungsversuche aufgezählt:

- theoretisches und praktisches Wissen (Thiel, 2007),
- explizites vs. implizites Wissen (Polanyi, 1966),
- semantisches und episodisches Wissen (Tulving, 1972),
- prozedurales Interpretationswissen (Reif, 1987),
- träges Wissen (Whitehead, 1929),
- strategisches Wissen (Alexander & Judy, 1988; Polya, 1957),
- situationsbedingtes Wissen (De Jong & Ferguson-Hessler, 1996),
- deklaratives, prozedurales (Anderson, 1976) und
- konditionales Wissen (Alexander & Judy, 1988).

Heute besteht relative Einigkeit darüber, wie Wissen strukturiert ist. Weitestgehend durchgesetzt hat sich die Unterteilung von deklarativem und prozeduralem Wissen (z. B.: De Jong & Ferguson-Hessler, 1996; Edelman, 2000; Lukesch, 2001; Rost et al., 2004; Schnotz, 1994; Seel, 2003; Steiner, 2001). Es bleibt aus kognitionspsychologischer Perspektive zu klären, wie diese beiden Wissensarten voneinander abzugrenzen sind. Des Weiteren besteht Klärungsbedarf bezüglich des Zusammenwirkens der beiden Wissensarten bei der Entstehung und Anwendung kognitiver Schemata (Schnotz, 1994). Daher soll das Augenmerk in der vorliegenden Arbeit auf die unterschiedlichen Auffassungen zur Struktur von deklarativem und prozeduralem Wissen gerichtet werden. Zu den prominentesten Theorien zu Wissensstrukturen, in denen diese Unterteilung aufgenommen wurde, zählt die in Abbildung 3 dargestellte ACT-Theorie (*Adaptive Control of Thoughts*) (Anderson, 1983). Anderson bezeichnet sein Werk nicht als eine allumfassende Theorie sondern vielmehr als einen ersten und ausbaufähigen Ansatz. Trotz dieser Einschränkung wird die Theorie oft zitiert, In ihr wird in die drei Gedächtnissysteme prozedurales und deklaratives Langzeitgedächtnis sowie das Arbeitsgedächtnis unterschieden.

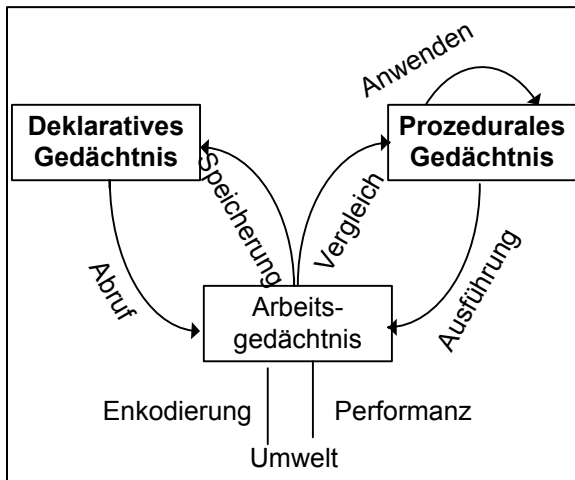


Abbildung 3: Das ACT-Modell von Anderson

Aus *Psychologie des Lernens und Lehrens* (S. 145) von H., Lukesch, 2001, Regensburg: Roderer Verlag. Copyright 2001 bei Roderer Verlag. Wiedergabe mit Genehmigung.

Amerkungen. Übersetzung v. Verf..

Das Arbeitsgedächtnis beinhaltet die in einer aktuellen Situation benötigten, verfügbaren Informationen aus dem deklarativen und prozeduralen Langzeitgedächtnis. Zunächst müssen dazu die situationsrelevanten Informationen dekodiert werden. Dann werden die relevanten Informationen aus dem deklarativen Gedächtnis abgerufen bzw. aus dem prozeduralen Gedächtnis ausgeführt. Das deklarative Wissen gelangt über Speicher- und Verstärkungsprozesse in das deklarative Gedächtnis und wird über Abfrageprozesse im Arbeitsgedächtnis aktiviert. Im Langzeitgedächtnis wird das Wissen in Form von Propositionen, Reihenfolgen und räumlichen Bildern abstrakt nach Sinngehalt sowie bereichsspezifisch abgespeichert. Die Elemente prozeduralen Wissens werden mit den situationsbestimmenden Bedingungen abgeglichen und im Ausführungsprozess passende Handlungen angewendet. Nach Anderson (2001) kann das deklarative Wissen, z. B. über die Ausführung eines naturwissenschaftlichen Experimentes, durch wiederholte Übung in Prozeduren umgewandelt werden und somit in das prozedurale Gedächtnis übergehen. In realen Situationen der Wissensanwendung interagieren deklarative Wissensanteile mit prozeduralen Wissensstrukturen im Arbeitsgedächtnis. Die Unterscheidung zwischen diesen beiden Wissensarten wird von der Mehrzahl der Wissenspsychologinnen und -psychologen getragen und hat in die einschlägigen Lehrbücher Eingang gefunden (z. B. Edelmann, 2000; Lukesch, 2001; Steiner, 2001).

Einige Wissenschaftlerinnen und Wissenschaftler unterscheiden auch zwischen dem Wissens- und Operatorgedächtnis, wobei deklaratives Wissen im Wissensgedächtnis (Prinz, 1983) und prozedurales Wissen im Operatorgedächtnis (Dörner, 1976) abgespeichert ist. Nach einer ersten Vorstellung der Theorien zur Struktur des Wissens, sollen die beiden herausgearbeiteten Wissensarten deklaratives und prozedurales Wissen näher beschrieben werden. Diese Beschreibung

dient der weiteren Untermauerung der bereits vorgenommenen Unterteilung von Kompetenz in Biologie in die Teilkompetenzen Konzept- und Prozesswissen. Anschließend wird die Übertragung dieser Unterteilung in die Naturwissenschaften vorangetrieben.

2.1.1 Deklaratives Wissen

Deklaratives Wissen ist das sprachliche, begriffliche und faktische Wissen, welches dem Individuum ermöglicht, die es umgebende Welt zu interpretieren. Schnotz (1994) bezeichnet deklaratives Wissen als „gespeicherte Datenstruktur über die Realität“ (Schnotz, 1994, S. 36). Dieses besteht sowohl aus singulärem Faktenwissen als auch aus allgemeinen Zusammenhängen zwischen diesen Fakten (Collins & Quillian, 1969; Schnotz, 1994). Es ist in hierarchisch organisierter Form gespeichert (Collins & Quillian, 1969) und kann in vielfältiger Form abgerufen werden. Deklaratives Wissen beinhaltet das „Wissen was“, also das gesamte sprachliche, begriffliche und faktische Wissen (Anderson, 1983). Auch das faktische Wissen um Methoden und Prozesse ist dieser Wissensart zuzuordnen (Schnotz, 1994). Es ist potenziell bewusheitsfähig – also explizit – und kann direkt beschrieben, über Verstehensprozesse unmittelbar erlernt und abgerufen werden (Schnotz, 1994). Für das Bilden der Begriffe des deklarativen Wissens wird der Schulbesuch als bedeutendster Faktor angesehen (Edelmann, 2000; PISA-Konsortium Deutschland, 2007).

Das deklarative Wissen ist in die Substrukturen episodisches und semantisches Gedächtnis unterteilt (Tulving, 1972). Das episodische Wissen umfasst alle individuellen (auto)biographischen Daten und Erfahrungen eines Individuums und ist räumlich und zeitlich gebunden. Dies kann auch Begebenheiten umfassen, von denen wir gehört oder gelesen haben. Das semantische Wissen basiert auf Erfahrungswissen, welches Personen über die Lebensspanne angesammelt haben (Edelmann, 2000; Tulving, 1972). Erlernte Begrifflichkeiten werden im deklarativen Gedächtnis als konzeptuelle Strukturen und mentale Modelle aufgebaut. Da semantisches Wissen in den fachspezifischen Kompetenzen als Bestandteil des Konzeptwissens angesehen werden kann, soll dieses näher erläutert werden.

Die kleinsten Informationseinheiten des semantischen Gedächtnisses, die Propositionen, sind hierarchisch angeordnet. So sind z. B. unter dem Oberbegriff Vogel weitere Unterbegriffe in Form von Exemplaren dieses Begriffs (z. B. Rotkehlchen) oder Eigenschaften dieser Exemplare (z. B. Federkleid, leichte, hohle Knochen oder beides) abgespeichert. Diese Begriffe auf den unterschiedlichen Hierarchieebenen sind über Bedeutungsrelationen wie trägt ein (Federkleid) oder hat (leichte, hohle Knochen) verbunden (Collins & Quillian, 1969; Seel, 2003; Steiner, 2001). Auch wenn hierzu noch keine einheitliche Theorie vorliegt (Edelmann, 2000), scheint in der Kognitionspsychologie die Idee zu überwiegen, dass diese Propositionen in Schemata abgebildet

werden, im beschriebenen Beispiel in dem Schema Vogel. Schemata sind umfassende, komplexe kognitive Wissensstrukturen, die als Erkenntnismittel bei der Erfassung der Umwelt dienen. In einem Schema werden Eigenschaften von Dingen in allgemeinen Kategorien über abstrakte begriffliche Teilsysteme oder konkrete Prototypen zusammengefasst. Diese bilden Grundbausteine für komplexere Netzwerke, die semantischen Netzwerke (Rumelhart & McClelland, 1986). Dieses hypothetische Konstrukt stellt Relationen zwischen spezifischen Zusammenhängen und begrifflichen Elementen unseres Denkens her (Steiner, 2001). Als Beispiel aus der Biologie kann hier das semantische Netzwerk zum Sachverhalt Wasser genannt werden. Wissen hierzu kann das Wissen um unterschiedliche Aggregatzustände, das Vorkommen in der natürlichen und vom Menschen beeinflussten Umwelt oder physikalische und chemische Eigenschaften beinhalten. Diese Netzwerke spielen für das Erkennen von Begriffen und das Lösen von realen Problemen eine immanente Rolle.

Wenn ein Individuum nun in einer Situation auf ein bestimmtes Problem trifft, werden mit dieser Situation im Zusammenhang stehende Oberbegriffe und weitere relevante Informationen und Strukturen bezogen auf diese Oberbegriffe aktualisiert. Diese Begriffe und Informationen werden dann in Form eines mentalen Modells im Arbeitsgedächtnis abgebildet (Gentner & Stevens, 1983). Mentale Modelle sind kognitive Konstruktionen, die die Realität möglichst plausibel über systematische Repräsentationen darstellen (Edelmann, 2000; Gentner & Stevens, 1983; Seel, 2003). Diese mentalen Modelle enthalten unterschiedliche Arten von Wissen und können sowohl deklarative als auch prozedurale Informationen enthalten. So sind zur Beantwortung der Frage, warum das Rotkehlchen fliegen kann, nicht die Farbe des Federkleides oder die Form des Schnabels sondern vielmehr der Knochenbau und die Flügel von Bedeutung. Die hier genannten Beispiele stehen für Objektschemata (Objekt Vogel, Objekt Wasser). Individuen können jedoch auch Handlungs- und Ereignisschemata abspeichern (Seel, 2003). Diese werden als Skripts bezeichnet und sollen im Zusammenhang mit prozeduralem Wissen näher erläutert werden.

2.1.2 Prozedurales Wissen

Das prozedurale Wissen wird auch als non-deklaratives Wissen bezeichnet, da die darunter subsumierten Fähigkeiten sehr heterogen sind (Lukesch, 2001). Hierbei handelt es sich um das „Wissen wie“, also um verfügbare Operationen oder geistige Prozesse, die Handlungen, Handlungsprogramme oder geistige Prozesse steuern, welche eine Person bezogen auf einen bestimmten Realitätsbereich ausführen kann und die nach einem Plan verkettet werden können. Seel merkt an, dass prozedurales Wissen Individuen ermöglicht, „komplexe kognitive Prozesse und/oder motorische Handlungen durchzuführen“ (Seel, 2003, S. 208). Der Begriff prozedurales Gedächtnis wurde erstmals von Ryle (1949) eingeführt, um die von ihm referierte Trennung von

Fakten- und Handlungswissen zu verdeutlichen. Dieses Handlungswissen beinhaltet neben den psychomotorischen Prozeduren auch kognitive Fertigkeiten wie die Ausführung einer Rechenaufgabe (Lukesch, 2001).

Der Erwerb prozeduralen Wissens findet über einen langsamen Prozess der schrittweisen Internalisierung und Automatisierung äußerer gegenständlicher Handlungen zu inneren geistigen Handlungen statt (Wygotski, 1964). Die Idee des prozeduralen Gedächtnisses setzt die Existenz eines Subgedächtnisses voraus, in welchem Operationen – also Handlungselemente – in definierten Spuren gespeichert werden. Diese Spuren enthalten Anweisungen, die eine Handlung in verschiedenen, variablen Situationen auslösen, wenn bestimmte äußere und innere Bedingungen vorliegen (Dörner, 1976). Die Reihenfolge des Abrufs ist stärker vorgegeben als beim deklarativen Wissen. Sie wird von einem Bedingungs- und Ausführungsgefüge – einer Wenn-Dann Beziehung – hervorgerufen (Schnotz, 1994; Seel, 2003). Der Wenn-Teil, also die äußeren oder inneren Bedingungen, beinhaltet deklaratives Wissen, der Dann-Teil die zu vollführende Prozedur (Seel, 2003). Bei der Repräsentation einer Handlung handelt es sich um einen „Eintrag in das semantische Gedächtnis, der eine reguläre Ereignisfolge in einer bestimmten Situation oder in einem bestimmten Kontext ... umfasst“ (Lukesch, 2001, S. 133). Hierbei spielen unbewusste kognitive Operationen eine Rolle, also geistige Prozesse, die in Form von sogenannten Produktionen repräsentiert sind (Opwis, 1988). Diese Produktionen können als Anweisungen bzw. Befehle für den Vollzug von Handlungen innerhalb eines bestimmten Realitätsbereichs verstanden werden (Arbinger, 1998). Die Handlung, die im Abgleichprozess die meisten Übereinstimmungen aufweist – also die plausibelste Handlung – wird ausgeführt. Handlungen werden demzufolge durch die Innensteuerung eines Subjektes und einer Entscheidung zwischen Handlungsalternativen hervorgerufen. Das Ergebnis einer Handlung ist eine Situationsänderung (Edelmann, 2000). Da die in diesem Zusammenhang durchgeführten Operationen unbewusst ablaufen und nur unmittelbar durch Übung mitzuteilen sind, wird das prozedurale Wissen oft als implizites Wissen (Polanyi, 1966; Seel, 2003) bezeichnet.

Die Ausführung einer Handlung führt zu Handlungsfolgen. Diese einzelnen Operationen verbinden sich bei wiederholten Anwendungen zu einem Handlungsschema, welches keine explizite Repräsentation des deklarativen Wissens voraussetzt und für vergleichbare Handlungen generalisiert wird. Bei der Zusammensetzung einzelner Handlungen zu Abläufen entstehen so genannte Skripts, die auch Ereignisschemata genannt werden (Anderson, 1989). Sie spezifizieren Handlungsabläufe wie ein Drehbuch die Handlung eines Filmes definiert und geben Personen in alltäglichen Situationen (wie z. B. dem Besuch einer Schulstunde) Orientierung (Lukesch, 2001). Diese Skripts können in variablen Situationen abgerufen und abgearbeitet werden. So wird eine Schülerin oder ein Schüler bei der erstmaligen Durchführung eines Experiments zum Pflanzenwachstum nach extern

vorgegebenen Anweisungen vorgehen. Dieses Handeln ist noch stark bewusstes Handlungswissen, das in seiner ursprünglichen Form sogar als deklarativ bezeichnet werden kann (VanLehn, 1989). Es kann nur mit Hilfe von externer Anleitung abgerufen werden. Wird der Untersuchungsplan in weiteren vergleichbaren Experimenten wiederholt, wird die Schülerin bzw. der Schüler souveräner und autonomer vorgehen, bis die Handlungsfolgen verinnerlicht sind. Nach weiteren Wiederholungen werden solche Schemata automatisiert und routinemäßig durchgeführt (wie z. B. Fahrradfahren oder Schlittschuhlaufen). Neue Handlungsabläufe und Produktionen werden also über einen Abgleich mit der Geschichte der ehemaligen Anwendungen verinnerlicht. Nach Dörner (1976) wird im Zusammenhang mit prozeduralem Wissen zwischen Primärprozessen – den automatisierten Prozessen – und Sekundärprozessen unterschieden. Bei letzteren ist das Wissen nur teilweise abrufbar und Individuen müssen gezielt nach geeigneten Operationen suchen. Ein diesbezügliches Beispiel beschreibt Anderson (2001) in seinen Ausführungen zum prozeduralen Gedächtnis. So wurde ein Computer-Experiment durchgeführt, in dem Probanden versuchen sollten, über die Regulation von Arbeitskräften die Produktionsmenge einer virtuellen Zuckerfabrik zu steuern (Berry & Broadbent, 1984). Dieses Computerexperiment ähnelt in vielen Merkmalen einem naturwissenschaftlichen Experiment.

Prozedurales Wissen umfasst folglich nicht nur automatisierte Prozesse, sondern auch Schritt-für-Schritt Prozesse in realen Situationen wie das bereits erwähnte Lösen einer Rechenaufgabe (De Jong & Ferguson-Hessler, 1996). Über den Begriff Problemlöseprozess werden Handlungen beschrieben, durch die in einem Problemraum nach Wegen zu einem definierten Ziel gesucht wird (Klieme et al., 2001). Bei dieser kognitiven Leistung wird ein Handlungsablauf nicht unterbewusst abgearbeitet. Vielmehr werden während des Handlungsablaufs bewusst Entscheidungen getroffen und Handlungen vorgenommen, um einen Zielzustand zu erreichen. Anderson (2001) geht daher davon aus, dass „alles prozedurale Wissen seinen Ursprung im Problemlösen besitzt“ (Anderson, 2001, S. 243). Die richtige Herangehensweise an ein Problem zeichnet sich also durch ein Zusammenspiel von begrifflichem und prozeduralem Wissen aus (Steiner, 2001). Eine kognitive Prozedur kann auch als eine Verkettung von Operationen (also Handlungen) nach einem Plan angesehen werden, der auf die Erreichung eines bestimmten Zielzustandes ausgerichtet ist (Seel, 2003). Diese Anwendungen müssen von dem Individuum in einem Problemraum erzeugt werden. Es werden erste Überschneidungen zu Problemlöseprozessen deutlich, die im weiteren Verlauf wieder aufgegriffen werden (s. Seite 65) und eng mit dem Zusammenspiel von prozeduralem und deklarativem Wissen zusammenhängen.

2.1.3 Das Zusammenspiel von prozeduralem und deklarativem Wissen

Es wurde gezeigt, dass prozedurales (in Abbildung 4 als reflexiv [nicht deklarativ]) bezeichnet) und deklaratives Wissen als voneinander abgrenzbare Wissensstrukturen konzeptualisiert werden können. Diese Abgrenzung der Wissensstrukturen wird auf Kompetenz in Biologie übertragen. Abbildung 4 verdeutlicht diese Wissensunterteilung erneut. Die Abbildung macht weiterhin deutlich, dass das deklarative Wissen wiederum in das semantische und episodische Wissen differenziert. Das semantische Wissen wird in der vorliegenden Arbeit weiter verfolgt, da es im Zusammenhang mit den Konzeptaspekten steht. Bezogen auf das prozedurale bzw. reflexive Wissen dient das in der Abbildung deklarierte Gedächtnis für Fertigkeiten in der fachspezifischen Ausprägung als Grundlage für das Prozesswissen.

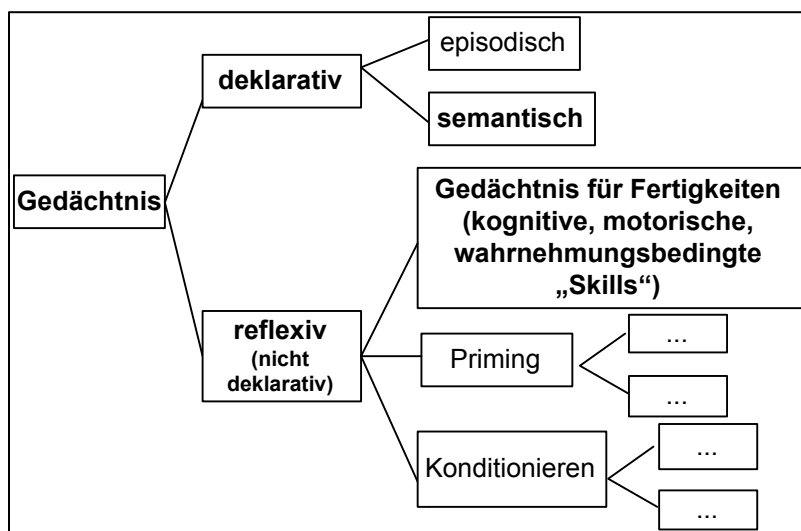


Abbildung 4: Darstellung der Wissensstrukturen von Markowitsch

Aus *Neuropsychologie des Gedächtnisses* (S. 7) von H. J. Markowitsch, 1992, Göttingen: Hogrefe. Copyright 1992 bei Hogrefe. Wiedergabe mit Genehmigung.

Anmerkungen. Die fettgedruckten Bezeichnungen verdeutlichen Inhalte der Wissensstrukturen, die im Zusammenhang mit der Untersuchung von Kompetenz in Biologie weiter verfolgt werden.

Psychologische Theorien zur Unterscheidung der Wissensarten – auch des deklarativen und prozeduralen Wissens – wurden bisher selten auf die Struktur disziplinspezifischen Wissens bzw. disziplinspezifischer Kompetenzen ausgeweitet. Zurückgreifend auf die Definitionsherleitung von Kompetenz in Biologie könnte sich das in der Pädagogischen Psychologie und Entwicklungspsychologie stark verwendete deklarative Wissen im Teilbereich Konzeptwissen wiederfinden. Prozedurales Wissen, welches noch nicht in die autonome Phase übergegangen ist, könnte der Prozesskomponente der genannten Definition zugeordnet werden.

In fachübergreifenden und -spezifischen Lernprozessen werden sowohl prozedurales als auch deklaratives Wissen abwechselnd benötigt. Anderson (1983) entwickelte in diesem Zusammenhang Stufen des Kompetenzerwerbs auf Grundlage seines ACT-Modells. Demnach kann deklaratives Handlungswissen in prozedurales Handlungswissen überführt werden. Andersons Theorie verdeutlicht die Wichtigkeit von elaboriertem deklarativem Wissen für den Erwerb und die Anwendung von Fertigkeiten und Prozeduren. Dies bedeutet aber auch, dass zunächst deklaratives Wissen erworben wird. Dieses wird nach zahlreichen Wiederholungen und Übungen in der assoziativen Phase der Wissenskompilierung in prozedurales Wissen überführt und in der autonomen Phase optimiert. Die Optimierung ist mit der Automatisierung gleichzusetzen (Ackerman, 1989). Weiterhin wird in Andersons ACT-Modell beschrieben (s. Abbildung 3), dass in einer realen Situation zur Bewältigung von Problemen beide Wissensformen im Arbeitsgedächtnis aktualisiert werden. So wird bei einer Aufgabenstellung, bei der Begegnung mit einem Problem oder beim Vorliegen einer Bedingung für einen Handlungsablauf zunächst das Faktenwissen für diese spezielle Situation abgerufen. In einem weiteren Schritt wird dieses Wissen mit den passenden Prozeduren in Verbindung gebracht.

In der naturwissenschaftsdidaktischen Forschung liegen bisher nur wenige Studien zum Zusammenspiel der beiden Wissensarten im Verlauf des Lernprozesses vor. Für den Grundschulbereich liegt eine Literaturstudie zur Entwicklung der Fähigkeiten des naturwissenschaftlichen Denkens vor, der dieses Zusammenspiel aufgreift (Zimmerman, 2007). Bezogen auf die beiden Wissensbereiche wird von einer gegenseitigen Beeinflussung der Entwicklung während der Durchführung einer Untersuchung ausgegangen. So scheint das Konzeptwissen innerhalb einer akademischen Disziplin die Auswahl von Hypothesen, Variablen und experimentellen Designs sowie die Bewertung von Ergebnissen zu beeinflussen. Durch die Erfahrungen aufgrund der Anwendung einer naturwissenschaftlichen Erkenntnismethode kann wiederum neues Konzeptwissen gewonnen werden (Zimmerman, 2007). Inadäquates Konzeptwissen müsste somit zwangsläufig zu Schwierigkeiten im Erlernen naturwissenschaftlicher Denkweisen und in der Entwicklung des Prozesswissens führen. Die gegenseitige Beeinflussung in der Entwicklung kann auch konterkarierende Folgen haben. So fällt es Individuen schwerer, neue Theorien aufgrund von (selbst gewonnenen) Untersuchungsergebnissen zu bilden, wenn diese mit ihren vorangegangenen Vorstellungen nicht übereinstimmen (Chinn & Brewer, 1998; Ruffman, Perner, Olson & Doherty, 1993; Zimmerman, 2007). Vielmehr versuchen Personen, die eigenen Theorien an Ergebnisse anzupassen (Zimmerman, 2007), die Daten zu ignorieren oder abzulehnen, sie als für nicht relevant in Bezug auf die Theorie zu deklarieren oder sie zurückzustellen bzw. neu zu interpretieren (Chinn & Brewer, 1998). Es ist auch möglich, dass Theorien peripher geändert werden oder die Daten akzeptiert und Theorien geändert werden (Chinn & Brewer 1998).

In der Forschung zu Wissensstrukturen und –erwerb in der Mathematik und dessen Entwicklung liegen mehrere Veröffentlichungen vor, die den Erwerb von konzeptuellem und prozeduralem Wissen untersuchen. Es wird argumentiert, dass Schülerinnen und Schüler sowohl Konzepte als auch Prozesse erlernen müssen, um mathematische Probleme lösen zu können. Des Weiteren scheint der Wissenserwerb ein iterativer Prozess von Konzept- und Prozesswissen zu sein (Schneider, Rittle-Johnson & Star, 2011).

Exkurs konzeptuelles und prozedurales Wissen in der Mathematik

Konzeptwissen in der Mathematik wird in den folgend beschriebenen Untersuchungen als explizites oder implizites Verständnis von maßgeblichen Prinzipien innerhalb der Disziplin und deren Relationen zueinander verstanden; als prozedurales Wissen werden die Aktionssequenzen zum Problemlösen angesehen. Diese beiden Wissensarten werden als Kontinuum verstanden. Für die Mathematik konnte gezeigt werden, dass diese beiden Wissensarten valide und reliabel gemessen werden können (Schneider et al., 2011; Schneider & Stern, 2010). Ähnliche Untersuchungen liegen in der Biologie nicht vor. Aufbauend auf ersten empirischen Grundlagen, die zeigten, dass

- Kinder mit besserem konzeptuellem Verständnis auch bessere Prozessfähigkeiten besitzen,
 - sowohl konzeptuelles Wissen den Prozessfähigkeiten vorausgeht als auch die Prozessfähigkeiten dem konzeptuellem Wissen,
 - Vermittlung von Konzepten als auch von Prozessen zu verbessertem Prozesswissen führen kann und
 - Verbesserung von Konzeptwissen zur Generierung von Prozessen führt,
- wurden Experimente zu spezifischen mathematischen Wissensseinheiten wie z. B. Gleichungen oder Dezimalzahlen durchgeführt.

In einem ersten Experiment beantworteten 86 Schülerinnen und Schüler der Jahrgangsstufen vier und fünf zwei Tests zum konzeptuellen Wissen (Rittle-Johnson & Wagner Alibali, 1999). Schülerinnen und Schülern, die falsche Antworten gaben, wurde ein Training zu konzeptuellem oder prozeduralem Wissen gegeben. Im Anschluss bearbeiteten alle Testteilnehmerinnen und Testteilnehmer einen Test zu konzeptuellem und prozeduralem Wissen. So konnte der Einfluss von neuem Wissen in der einen Wissensart auf die jeweils andere Wissensart untersucht werden. Es zeigte sich, dass der Aufbau von Wissen in einer Wissensart auch zu einer Verbesserung in der jeweils anderen Wissensart führen kann. Das Lernen von Konzepten hatte einen stärkeren Einfluss auf das Prozesswissen als das explizite Lernen spezifischer Prozesse. Schülerinnen und Schüler, die Konzeptwissen vermittelt bekamen, wendeten unterschiedliche Prozesse zur Problemlösung an und konnten diese auch besser auf verwandte Probleme transferieren.

In einer weiteren Untersuchung konnte gezeigt werden, dass zu Beginn einer Intervention ein besseres Prozesswissen mit einem besseren Konzeptwissen einherging ($r = .33$) (Rittle-Johnson, Siegler & Wagner Alibali, 2001). In einem experimentellen Design wurden in einem Pretest 49 Schülerinnen und Schüler mit unterschiedlichem Vorwissen über Dezimalbrüche am Ende der fünften Klasse getestet. Danach bearbeiteten die Kinder eine Intervention, die das Lernen von Dezimalzahlen beinhaltete. Mit einem anschließenden Posttest wurde untersucht, ob das Vorwissen der Kinder auch zu einem verbesserten Prozesswissen führte. Durch die Intervention konnte sowohl das Prozesswissen als auch das Konzeptwissen verzeichnet werden.

Auch Perry (1991) konnte zeigen, dass die kombinierte Vermittlung von Prozess- und Konzeptwissen ähnliche Effekte auf den Problemtransfer hatte wie die alleinige Vermittlung von Prozesswissen. Die beste Transferleistung erbrachten Schülerinnen und Schüler, die nur Konzeptwissen vermittelt bekamen. Mögliche Gründe für diese gegenseitige Beeinflussung der Wissensarten könnten sein:

- die Schülerin bzw. der Schüler gewinnt die Einsicht, dass Problemlöseprozesse mit vorhandenem Wissen (z. B. mit Fehlkonzepten) inkonsistent sind bzw. verinnerlicht Vorwissen nach dessen erfolgreicher Anwendung,
- Konzeptwissen bedingt die Generierung von Prozessen, da es essenzielle Elemente von korrekten Problemlöseprozessen darstellt,
- über Selbsterklärung richtiger Prozeduren – und über die Reflexion dieser – wird das Konzeptwissen verändert und
- durch den Zuwachs an Prozesswissen werden Ressourcen freigesetzt, die für die Akquise von Konzeptwissen verwendet werden können (Rittle-Johnson & Wagner Alibali, 1999).

Somit könnte innerhalb des Problemlöseprozesses das Konzeptwissen den Schülerinnen und Schüler helfen, Hauptmerkmale der anzuwendenden Prozesse zu entschlüsseln.

Auf Grundlage ihrer experimentellen Untersuchungen stellten Rittle-Johnson et al. (2001) ein Modell vom iterativen Prozess der Entwicklung von konzeptuellem und prozeduralem Wissen für die Mathematik auf. Dieses ist in Abbildung 5 dargestellt.

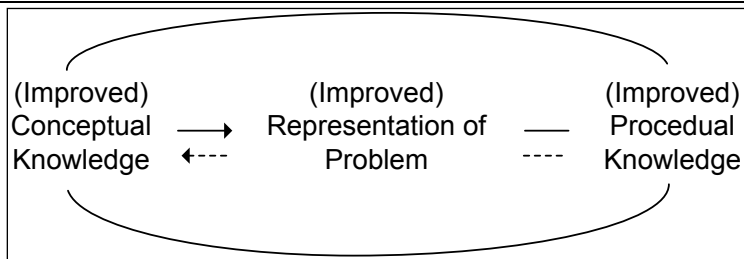


Abbildung 5: Iteratives Modell des Erlernens von Konzept- und Prozesswissen

Aus „Developing Conceptual Understanding and Procedural Skill in Mathematics: An Iterative Process“, von B. Rittle-Johnson et al., 2001, *Journal of Educational Psychology*, 93, S. 347. Copyright 2001 bei American Psychological Association. Wiedergabe mit Genehmigung von Bethany Rittle-Johnson.

Das abgebildete Modell verdeutlicht die wechselseitige Beeinflussung und erneut die Überlappung der beiden Kompetenzaspekte. Rittle-Johnson und Wagner Alibali (1999) schlagen vor, die Ergebnisse auf weitere Disziplinen auszuweiten und in einem ökologisch valideren Umfeld, wie z. B. dem Klassenraum, sowie an einer heterogeneren Stichprobe zu testen.

Auch wenn eine detaillierte Untersuchung des Erwerbs von Konzept- und Prozesswissen in der vorliegenden Arbeit nicht angestrebt wird, so geben die referierten Ergebnisse doch interessante Hinweise auf einen Zusammenhang zwischen den distinkten Teilbereichen von Kompetenz. Im Lichte von Andersons ACT-Modell sind die Ergebnisse zweideutig. Anderson (2001) postuliert, dass zunächst deklaratives Wissen aufgebaut wird, welches in einem späteren Stadium prozeduralisiert wird (s. Seite 51). Diese Annahme konnte mit den fachspezifischen Untersuchungen in Mathematik nur teilweise bestätigt werden. Konzepte scheinen beim Lösen von Problemen und beim Erwerb von Prozesswissen im Vordergrund zu stehen. In einer Problemlösesituation werden beide Wissensarten aktualisiert. Auch Alexander und Judy (1988) konnten in ihrer Literaturstudie zeigen, dass strategisches, disziplinspezifisches Wissen zur Anwendung und zum Erwerb von disziplinspezifischem Fachwissen beiträgt. Außerdem verfügten Schülerinnen und Schüler mit fehlendem Inhaltswissen nur über geringes strategisches Wissen.

Ähnlich detaillierte Untersuchungen zur Unterscheidung von Konzept- und Prozessaspekten von Kompetenz und deren Beziehung zueinander fehlen für die Naturwissenschaften. Die Bestrebungen der einzelnen Fachdidaktiken richten sich vielmehr auf die Ausdifferenzierung von Teilaspekten wie der Modellkompetenz (Upmeyer zu Belzen & Krüger, 2010), der Bewertung (Eggert & Bögeholz, 2006) oder der Erkenntnisgewinnung (Mayer, Grube & Möller, 2008), um eine Implementation von Kompetenzmodellen in den Unterricht zu erwirken. So werden für die Modellkompetenz die beiden Bereiche Kenntnisse über Modelle und Bildung von Modellen herausgearbeitet. Diese Unterteilung

wird auf der Unterscheidung zwischen deklarativem und prozeduralem Wissen vorgenommen (Upmeier zu Belzen & Krüger, 2010). Einzig der Artikel von Klos et al. (2008), und die Konzeptualisierung der Bildungsstandards (s. Seite 34) weisen auf eine gleichgeartete Unterscheidung der Wissens- und Kompetenzstrukturen in den Naturwissenschaften hin. Um beide Wissensarten für die Naturwissenschaften untersuchen zu können, gilt es zunächst, das Prozesswissen in den Naturwissenschaften und in der Biologie näher zu spezifizieren. Zwei Ansätze sind bei der Begriffsbestimmung von Erkenntnisgewinnung möglich.

2.1.4 Erkenntnisgewinnung in den Naturwissenschaften und in der Biologie

Unter Erkenntnisgewinnung in den Naturwissenschaften wird die Anwendung einer Fülle von Herangehensweisen, Methoden und Prinzipien subsumiert (Roberts, 2001; Zimmerman, 2007), die verwendet werden können, um ein naturwissenschaftliches Problem zu lösen (Nehring, Nowak, Tiemann & Upmeier zu Belzen, 2011). Fähigkeiten und Fertigkeiten im Zusammenhang mit Erkenntnisgewinnung in den Naturwissenschaften und der Biologie bergen neben dem unmittelbaren handlungsrelevanten Nutzen detaillierte Einsichten in aktuelle und zukünftige gesellschaftliche Themen. Zu diesen Themen wie z. B. der Stammzellforschung kann eine Person nur dann eine eigene und naturwissenschaftlich fundierte Meinung haben, wenn sie nicht nur über Fachwissen zu diesem Thema sondern auch über Wissen bezüglich Mitteln und Herangehensweisen in der Stammzellforschung verfügt. Dieses Beispiel verdeutlicht, wie wichtig eine Bildung über den Schulkontext und das Faktenwissen hinaus ist. Forschung in der Biologie und Biotechnologie ist von einer rasanten Entwicklung gekennzeichnet (Hurd, 1998). Zum einen sollte im Sinne der *scientific literacy* dieser Fortschritt Eingang in schulische Themen und die Unterrichtsgestaltung finden. Zum anderen kann im schulischen Biologieunterricht schwerlich jenes Fachwissen vermittelt werden, welches in 20 Jahren von Bedeutung sein wird. Im Biologieunterricht heutiger Erwachsener wurde mit großer Wahrscheinlichkeit die Bedeutung der Stammzellen in der Medizin nicht behandelt. Nur wenn sich Bürgerinnen und Bürger selbstständig Wissen über (biologische) Thematiken und Methodiken aneignen können, werden sie zu diesen eine Meinung haben und eine durchdachte Entscheidung treffen können. Shamos (1995) vertritt die Meinung, dass es den Bürgerinnen und Bürgern gar nicht möglich ist, sich das Wissen in dem nötigen Umfang anzueignen. Daher sollen sie im Unterricht eher ein Bewusstsein für die Besonderheiten der Disziplin entwickeln und ein Gefühl dafür, welche Informationen sie aus welchen Quellen beziehen können, um Entscheidungen zu treffen.

Unter dem Begriff Erkenntnisgewinnung im naturwissenschaftlichen Kontext – im deutschsprachigen Raum auch als wissenschaftliche Denk- und Arbeitsweisen oder

Wissenschaftspropädeutik bezeichnet – werden die Fähigkeiten naturwissenschaftliche Erkenntnisse gewinnen, naturwissenschaftliche Methodik anwenden sowie Einschätzen von Methodik und Aussagen zusammengefasst (Mayer, 2007). Im internationalen Diskurs wird der Begriff Erkenntnisgewinnung unter den Überbegriffen *scientific reasoning* oder *scientific inquiry* (Klahr, 2000; Möller, Grube & Mayer, 2007) aber auch *procedural understanding* (Roberts, 2001) subsumiert. Nicht nur die Begrifflichkeiten weisen große Unterschiede auf, auch die Konzeptualisierungen und Definitionen von Erkenntnisgewinnung sind vielfältig (z. B.: Klahr, 2000; KMK, 2005a; Mayer, 2007; Mayer et al., 2008; Roberts, 2001). Diese unterscheiden sich zunächst in der Weite des Konzepts, also in der inneren Ausgestaltung durch Komponenten. So konzeptualisieren einige Autoren Erkenntnisgewinnung ausschließlich als Wissen über die Forschungsmethoden in der Biologie (Roberts, 2001; KMK, 2005a). Andere beziehen manuelle Fähigkeiten oder Wissen über das naturwissenschaftliche Forschen ein (Gott, Duggan & Johnson, 1999; Klahr, 2000; Mayer, 2007).

Eine eingehende Betrachtung der unterschiedlichen Konzeptualisierungen ist an dieser Stelle nicht zielführend. Von Bedeutung im Kontext der vorliegenden Arbeit ist, dass im Bereich der Erkenntnisgewinnung Defizite im schulischen und beruflichen Kontext gefunden werden konnten (Baumert et al., 1997; Fischer et al., 2005; Gott et al., 1999; Grube, Möller & Mayer, 2007; Klieme & Leutner, 2006; Layton, 1973; Möller, Hartmann & Mayer, 2010; Möller, Hartmann & Mayer, 2009; Tamir, Stavy & Ratner, 1998). So zeigten Unterrichtsanalysen auf Grundlage der PISA-Studie 2006, dass das selbstständige Forschen und das Modellieren und Anwenden von naturwissenschaftlichen Konzepten selten in deutschen Unterrichtsräumen vorzufinden ist (Seidel et al., 2007). Zudem scheint das unterrichtliche Experimentieren oder der Einsatz von Modellen in den naturwissenschaftlichen Fächern eher der Förderung des Fachwissens als des Prozesswissens in den Naturwissenschaften zu dienen (Duit et al, 2001; Höttecke, 2001). Gelernt wird von Schülerinnen und Schülern jedoch nur, was auch aktiv gelehrt wird (Weinert, 2001). Daher wird dieser Aspekt der naturwissenschaftlichen Fächer verstärkt im Unterricht eingefordert (Möller et al., 2007; Shen, 1975). Neben den konkreten, direkten Erfahrungen sollen die Schülerinnen und Schüler die Methoden der Naturwissenschaften modellieren und erlangen ein Verständnis über die Naturwissenschaften. Damit geht eine Wertschätzung sowie die Unterstützung einer positiven Einstellung gegenüber den Naturwissenschaften einher (Tamir, Doran & Chye, 1992). Die TIMS-Studie zeigte allerdings, dass Schülerinnen und Schüler sowohl im konzeptuellen Verständnis als auch im Verständnis naturwissenschaftlicher Arbeitsweisen Defizite aufweisen (Baumert et al., 1997). Ein gesichertes empirisches Fundament über Kompetenzen und Wissenserwerb bezüglich der Inhalte und Prozesse der Erkenntnisgewinnung liegt noch nicht vor. Im deutschsprachigen Raum liegen erste Untersuchungen für das Fach Biologie vor (Mayer et al., 2008).

Unter Rückgriff auf die vorangegangenen referierte Struktur von Wissen aus der Kognitionspsychologie, kann Erkenntnisgewinnung im biologischen Kontext aus zwei verschiedenen Perspektiven betrachtet werden. Zum einen ist eine Auffassung als fachspezifisches prozedurales Wissen (Mayer et al., 2008; Hodson, 1992), also Anwendungswissen, möglich. Zum anderen beschreiben erste Konzeptualisierungen in der Biologie und den Naturwissenschaften Erkenntnisgewinnung als disziplinspezifische Problemlösekompetenz (Bayrhuber et al., 2007; Helgeson, 1994; Klahr, 2000; Mayer, 2007). Das Konstrukt Erkenntnisgewinnung wird daher im Folgenden aus diesen zwei Perspektiven beschrieben.

2.1.4.1 Erkenntnisgewinnung als biologiespezifisches prozedurales Wissen

Der prozedurale Aspekt biologischen Wissens wird im Alltag vorrangig als praktische Fähigkeit denn als tatsächliche Wissensbasis angesehen und selbst von praktizierenden Naturwissenschaftlerinnen und Naturwissenschaftlern nicht als genuin biologisches Wissen sondern als gesunder Menschenverstand gesehen (Shamos, 2002), der nicht mit den Naturwissenschaften bzw. der Biologie in Verbindung steht (Gott et al., 1999; Roberts & Gott, 1999, 2000; Wong & Hodson, 2009). Die Grundannahme einer Konzeptualisierung von Erkenntnisgewinnung ist jedoch die Existenz eines Wissensfundus, der das Verständnis naturwissenschaftlicher Beweisführung und naturwissenschaftlichen Arbeitens unterstützt (Gott et al., 1999). In der naturwissenschaftsdidaktischen Forschung werden *procedural understanding* und *concepts of evidence* näher untersucht; in der naturwissenschaftlichen Diagnostik stehen *process skills* sowie *scientific inquiry* im Vordergrund. In seinen Abhandlungen zu den Wissensstrukturen erwähnt Anderson (2001), dass Problemlösen der Ursprung des prozeduralen Wissens sei (s. Seite 55). Er deutet damit auf die schon eingeführten Sekundärprozesse hin; Prozesse, die noch stark von deklarativem Wissen bestimmt sind. Wenn disziplinspezifisches, prozedurales Wissen verortet werden soll, stehen nicht die automatisierten Prozesse in Vordergrund, die nach einem stark vorbestimmten, unterbewussten und nur wenig beeinflussten Skript ablaufen. Vielmehr muss es sich um die erwähnten Sekundärprozesse handeln. Dieses Wissen wird auch als *thinking beyond the doing* bezeichnet (Roberts, 2001). Aus kognitionspsychologischer Sichtweise bedarf es zur Ausführung (disziplinspezifischer) Prozeduren der folgenden drei Prozesse:

- Auswahl und Abruf eines Skripts,
- Anpassung der Strukturen des Skripts an die aktuelle Situation und
- Ausführung der Lösungsprozedur (Seel, 2003).

Um das relevante Skript auszuwählen, müssen vom Individuum Analogien zwischen dem aktuellen Problem und vorangegangenen Problemlösungen gefunden und erfolgreich übertragen werden

(Anderson, 2001). Die Schülerin bzw. der Schüler sollte also z. B. beim Sammeln von Pilzen⁴ mittels vorangegangenen angewandter Strategien aus der naturwissenschaftlichen Erkenntnisgewinnung ein gefundenes Exemplar identifizieren können. Die Schülerin bzw. der Schüler muss also eine zielgerichtete Strategie anwenden, um die einem bestimmten Inhaltsbereich zugehörige Aufgabe zu lösen (Alexander & Judy, 1988). Die erfolgreich angewendete Abfolge von Prozeduren wird dann als Produktionssystem abgespeichert (Arbinger, 1998; Lukesch, 2001) und kann in ähnlichen Problemsituationen abgerufen und angewendet werden. Als idealtypischer Verlauf einer Prozedur kann das Definieren eines Zieles, die Überprüfung der Anwendbarkeit vorhandener Regeln und das Ausführen der Aktion angesehen werden (Anderson, 2001). Beim Beispiel des Pilzesammelns wird das Ziel der Identifizierung eines gefundenen Pilzes über adäquate, oft mehrstufige Abgleichprozesse bestimmter Merkmale des Pilzes erreicht, z. B. mit Beschreibungen der Merkmale mithilfe eines Bestimmungsbuchs.

Es wird deutlich, dass auch die Ausführung disziplinspezifischer Prozeduren ein stark kognitiv geprägter Prozess ist. Der Prozess der Erkenntnisgewinnung in Biologie kann in die drei Teilbereiche (a) manuelle Fertigkeiten, (b) wissenschaftliches Denken und (c) Wissenschaftsverständnis (Mayer, 2007, s. Abbildung 6, S. 67) oder in die beiden Teilbereiche Fertigkeiten und Konzepte des Prozesswissens (Gott et al., 1999) unterteilt werden. Erkenntnisgewinnung konzeptualisiert als prozedurales Wissen würde nur die beiden Konstrukte manuelle Fertigkeiten und das wissenschaftliche Denken beinhalten. Das Wissensverständnis beschreibt Verstehens- und Verständnis- keine Handlungsprozesse. Manuelle Fertigkeiten und wissenschaftliches Denken sind durch grundlegende Elemente einer naturwissenschaftlichen Herangehensweise an Probleme gekennzeichnet, wie z. B. das Untersuchungsdesign oder die Erhebung von Daten und deren Präsentation, die Datenanalyse und Interpretation von Ergebnissen (Mayer, 2007). Wie auch das Problemlösen kann diese Abfolge nur eine idealisierte sein (Klos et al., 2008). Das Wissenschaftsverständnis kann eher dem Teilbereich Konzeptwissen biologischer Kompetenz und somit dem deklarativen Wissen zugerechnet werden. Innerhalb der Biologie können verschiedene Herangehensweisen angewendet werden: stark kontrollierte, laborbasierte Untersuchungen, schwer kontrollierbare Felduntersuchungen oder ökologische Gutachten (Roberts, 2001). Diese unterschiedlichen biologischen Herangehensweisen können wiederum in die Konzepte Instrumente, einzelne Messung, Untersuchungsdesign, Zusammenhänge, Daten von anderen und Validität/Reliabilität subsumiert werden (Roberts, 2001). Am Beispiel des Untersuchungsdesigns – welches auch am häufigsten im Unterricht thematisiert wird (Roberts & Gott, 2000) – bedeutet dies, dass die

⁴ Die alltägliche Formulierung „Pilze sammeln“ ist biologisch nicht korrekt. Der Pilz ist ein unterirdisch weit verzweigtes organisches System; an der Oberfläche befinden sich nur die Fruchtkörper dieses Systems. Biologisch einwandfrei wäre es also, vom Sammeln dieser Fruchtkörper zu reden. Um die Lesbarkeit zu gewährleisten, wird in dieser Arbeit die alltägliche Formulierung Pilze sammeln verwendet.

Schülerin bzw. der Schüler ein Verständnis der Struktur und der einzelnen Bestandteile einer Untersuchung hat. So sollte sie bzw. er z. B. wissen, was eine unabhängige Variable ist, wie und warum sie ausgewählt und manipuliert werden kann oder wozu eine Kontrollgruppe oder eine Messwiederholung dient und dieses Wissen in naturwissenschaftlichen Untersuchungen anwenden können. Erkenntnisgewinnung als biologiespezifisches Problemlösen umfasst jedoch weitere Aspekte, die im Folgenden vorgestellt werden.

2.1.4.2 Erkenntnisgewinnung als biologiespezifisches Problemlösen

Problemlösen ist zunächst die Überwindung der Diskrepanz zwischen einem Ausgangszustand und einem anvisierten Endzustand mittels logischer Operationen (Dörner, 1976; Funke, 2003; Mayer, 2007; Nehring et al., 2011; Newell & Simon, 1972). Hierbei wird zwar das vorhandene Wissen einer Person angewendet, es reicht jedoch nicht aus, um zum Endzustand zu gelangen. Das Wissen muss vielmehr neu konstruiert werden, um diesen Endzustand zu erreichen (Dörner, 1976). Routinierte Vorgehensweisen führen nicht zu diesem Endzustand (Klieme et al., 2001). So muss für eine effektive Problemlösung eine Vernetzung zwischen dem deklarativen und prozeduralen Wissen vorhanden sein (Edelmann, 2002). Die Problemlöserin bzw. der Problemlöser muss kognitive Prozess nutzen, um sich mit realen Problemstellungen auseinanderzusetzen und sie zu lösen (Koppelt & Tiemann, 2008). Problemlösen zeichnet sich durch sichtbare Tätigkeiten und Entscheidungen aus und wird in einzelnen Schritten realisiert. Problemlösen kann als „zielorientiertes Denken und Handeln in Situationen“ bezeichnet werden, „für deren Bewältigung keine routinierten Vorgehensweisen verfügbar sind“ (Klieme et al., 2001, S. 185). Konstitutiv für den Prozess des Problemlösens – also die Zielerreichung – ist das „Verstehen der Problemsituation und deren schrittweise Veränderung, gestützt auf planvolles und schlussfolgerndes Denken“ (ebd., S. 185).

Es wird deutlich, dass vom Problemlöser adäquate Prozeduren und Operationen zur Erreichung des Wunschzustandes eingesetzt werden müssen. Es bedarf also sowohl einer systematischen Abfolge von richtigen Prozeduren als auch eines Zusammenspiels von begrifflichem und prozeduralem Wissen (OECD, 2000; Zimmerman, 2007) sowie Fähigkeiten zum schlussfolgerndem Denken (Klieme et al., 2001). Die OECD hat im Zusammenhang mit der Testung fachübergreifender Problemlösekompetenz in der PISA-Studie 2003 folgende Schritte des Problemlöseprozesses als Strukturierung zu Grunde gelegt: (a) Problem verstehen, (b) charakterisieren, (c) repräsentieren, (d) lösen, (e) Problemlösung reflektieren und (f) kommunizieren (OECD, 2003). Diese wurden in weiterführenden Untersuchungen in der Chemiedidaktik zu vier Schritten zusammengefasst. Die Schritte Problem verstehen und charakterisieren sowie Lösung reflektieren und kommunizieren ließen sich in diesen Studien nicht trennen (Tiemann, Koppelt & Nehring, 2011). Edelmann (2000) unterteilt den kognitiven Prozess des Problemlösens in die Einzelbestandteile Problemraum,

Situationsanalyse, Suchraum, Lösung und Evaluation. Klahr (2000) erweitert Erkenntnisprozesse im Sinne von spezifischem naturwissenschaftlichem Problemlösen in seiner *scientific discovery as dual search* (SDDS)-Theorie um die zwei Suchräume Hypothesensuchraum (*space of hypotheses*) und den Experimentiersuchraum (*space of experiments*). Dieses Modell soll dazu dienen, allgemeingültige Aussagen über alle Verfahren der naturwissenschaftlichen Erkenntnisprozesse zu beschreiben (Hammann, 2007). Der Problemlöseprozess ist durch zwei ineinandergreifende Räume gekennzeichnet: Die Suche im Hypothesenraum und die Suche im Experimentierraum (Klahr, 2000). Um ein Phänomen zu erklären, werden zunächst Hypothesen gesucht und getestet. Im nächsten Schritt werden die Ergebnisse für und gegen die vorliegenden Hypothesen beschrieben. Diese Ergebnisse werden aus dem Experimentierraum herangezogen (Hammann, 2007). Nach den Schlussfolgerungen aus den Daten werden Entscheidungen gegen oder für Hypothesen gefällt. Diese werden weiter in verschiedene Hauptkomponenten unterteilt. So besteht das Schlussfolgern aus (a) den Daten, (b) der Prüfung der Ergebnisse und (c) der Entscheidung zur Annahme, zur Ablehnung oder zur weiteren Untersuchung des Problems. Die drei Schritte des Problemlösens sind durchaus umstritten, da sie einen idealtypischen und nicht den realen Vorgang des Problemlösens und der Erkenntnisgewinnung nachzeichnen (Lederman, Abd-El-Khalick, Bell & Schwartz, 2002). Auf Grundlage der SDDS-Theorie wurden zwei typische Herangehensweisen an Probleme identifiziert (Hammann, 2007). Die Experimentatoren testen eine im Vorhinein aufgestellte Hypothese. Wenn dieser Test negativ ausfällt, werden ohne neue explizit formulierte Hypothesen weitere Experimente durchgeführt. Die Theoretiker beginnen mit dem gleichen Prinzip. Wenn die erste Hypothesentestung nicht zum gewünschten Erfolg führt, formulieren sie explizit neue Hypothesen und führen nach diesen ihre konsekutiven Experimente durch. Sie kommen daher mit weniger Experimenten zum gewünschten Zielzustand. Das SDDS-Modell wurde im biologiedidaktischen Kontext bereits angewandt (Dunbar, 1993).

Der Weg zum Zielzustand muss vom Problemlöser also in zu bewältigende Teilprozesse unterteilt werden, die dann systematisch zu Produktionssystemen verkettet und abgearbeitet werden. Daher wird Problemlösen auch als eine Sonderform des planvollen Handelns bezeichnet (Edelmann, 2000). Die Qualität eines Problemlöseprozesses wird durch das Verständnis der Problemlösesituation, die Denkprozesse bei der Problembearbeitung und die Angemessenheit der erreichten Lösung bestimmt (Klieme et al., 2001). Der aktuelle Stand der Forschung zeigt, dass eine bereichsübergreifende Problemlösekompetenz nicht messbar zu sein scheint (Klieme et al., 2001). Problemlösefähigkeit sollte disziplinbezogen also z. B. als Problemlösen in den Naturwissenschaften oder in der Biologie erhoben werden (Klieme et al., 2001).

Verschiedene Autoren haben Erkenntnisgewinnung als Problemlöseprozess konzeptualisiert (Grube et al., 2007, Hammann et al., 2007; Klos et al., 2008; Mayer et al., 2008; Möller, Grube, Hartmann & Mayer, 2009). Mayer (2007) beschreibt Erkenntnisgewinnung als einen wissenschaftlichen, relativ komplexen, kognitiven und wissensbasierten Problemlöseprozess. Erkenntnisgewinnung bedeutet bei ihm das Anwenden von Wissen und Fähigkeiten in bestimmten Situationen und besteht aus den drei Dimensionen praktische Arbeitstechniken, wissenschaftliche Erkenntnismethoden und Charakteristika der Naturwissenschaften.

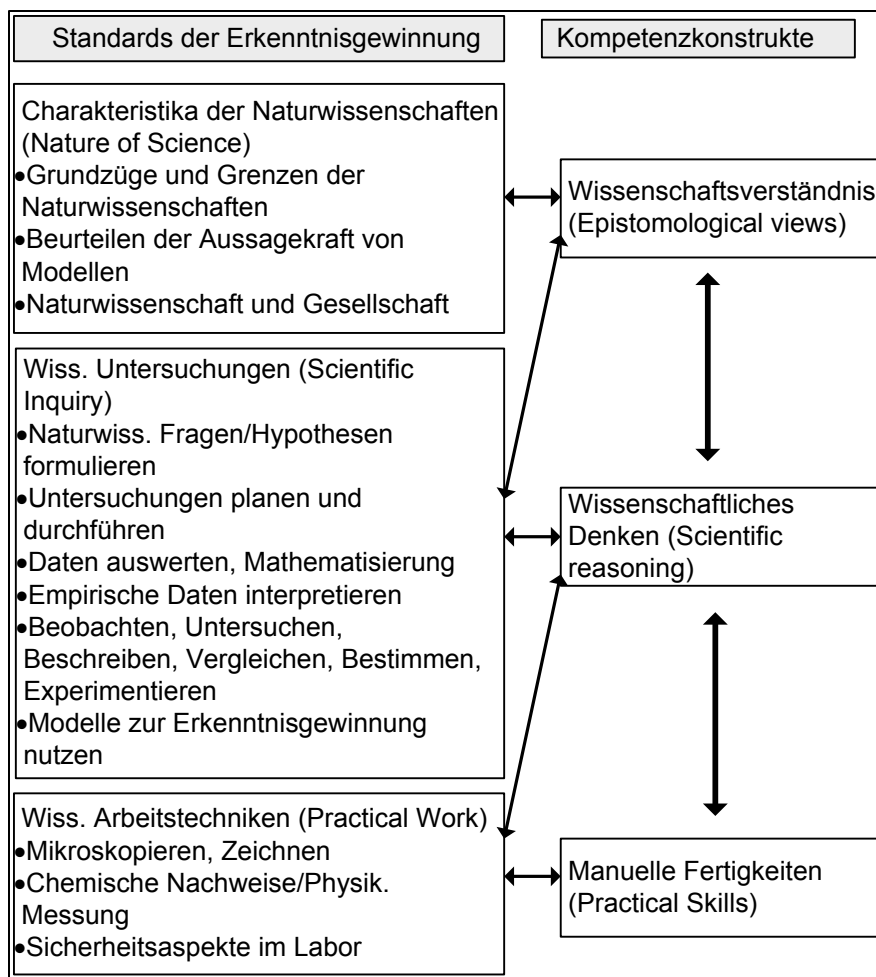


Abbildung 6: Rahmenkonzept wissenschaftsmethodischer Kompetenzen von Mayer (2007)

Aus *Erkenntnisgewinnung als wissenschaftliches Problemlösen* (S. 178) von J. Mayer, 2007. In: Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden, von Krüger, D. und Vogt, H. (Hrsg.), Berlin, Springer, S. 177-186. Copyright 2007 bei Springer-Verlag. Wiedergabe mit Genehmigung.

In Abbildung 6 werden diese drei Aspekte inhaltlich gefüllt. So beinhaltet das Wissenschaftsverständnis die abstrakte Vorstellung darüber, wie Forschung in den

Naturwissenschaften betrieben wird. Hierzu gehört das Wissen um die Grenzen der naturwissenschaftlichen Forschung, deren Beitrag für die Gesellschaft und das Bewerten und Einordnen von Produkten naturwissenschaftlicher Forschung. Ein konkretes Beispiel für diesen Aspekt ist das Wissen darüber, dass neu aufgestellte Theorien keine falschen Theorien ersetzen. Vielmehr soll von der Schülerin bzw. dem Schüler verstanden werden, dass überholte Theorien den Wissensstand zu einem bestimmten Zeitpunkt widerspiegeln. Das wissenschaftliche Denken umfasst die Methoden der naturwissenschaftlichen Forschung und deren Teilschritte. Die manuellen Fähigkeiten schließlich setzen sich aus dem Wissen über die praktischen Anwendungen zusammen.

Die Spezifität eines Problemlöseprozesses wird durch den Kontext (z. B. das Ökosystem Wald) und die notwendigen Prozeduren und Operationen (z. B. kriteriengeleitetes Vergleichen von Pflanzen) bestimmt. Anspruchsvolle disziplinbezogene Problemlöseprozesse setzen einen deklarativen und prozeduralen Wissensfundus in der jeweiligen Disziplin sowie Metakognition und kognitive Fähigkeiten voraus (Mayer, 2007). Ein Beispiel für die Operationalisierung von Problemlösekompetenz in den Naturwissenschaften ist das Virtuelle Labor (Klieme et al., 2001). In diesem virtuellen Labor (Autoren Reimann/Schmitt) sollen Schülerinnen und Schüler Gesetzmäßigkeiten induktiv aus einer Serie von selbst gestalteten, aufeinander aufbauenden Experimenten erschließen. Hierzu können sie eigene Hypothesen aufstellen, Bedingungen für Versuchsdurchführungen bestimmen oder variieren und so eine Serie von Experimenten durchführen. Die erstellten Versuche werden ihnen in einem Videoclip vorgeführt. Bei Erkenntnisgewinnung im Sinne eines spezifisch naturwissenschaftlichen (und biologischen) Problemlösens werden wie im beschriebenen Beispiel Strategien angewendet. Problemlösen durch Anwendung von geplanten Strategien liegt besonders bei Denkprozessen vor, die in äußerlich sichtbare Tätigkeiten und Entscheidungen erfolgen. Hierzu muss eine abstrakte Version des Problemraums abgebildet werden. In diesem wird das auftretende Problem abgebildet. Das abstrakte Problem wird sodann unter Anwendung einer Methode oder Strategie gelöst. Diese Lösung des abstrakten Problems wird als Plan zur Lösung des realen Problems herangezogen. Abschließend wird der Plan auf das Originalproblem angewendet, also ausgeführt (Klahr, 2000). Diese Problemlösestrategie wird auch als Mittel-Ziel-Analyse bezeichnet (Hammann, 2007). Deutsche Schülerinnen und Schüler scheinen im naturwissenschaftlichen Problemlöseprozess eher zur Anwendung von Versuch und Irrtum zu neigen (Mayer, 2007). Es können jedoch weitere Problemlösestrategien von Schülerinnen und Schülern angewendet werden. Beim sogenannten *hill climbing* werden mehrere Versuche in unterschiedliche Richtung ausprobiert. Danach wird der erfolgversprechendste Weg (der mit dem höchsten Anstieg) eingeschlagen. Bei der *means-end analysis* wird die Differenz zwischen dem Anfangs- und Zielzustand analysiert und dann nach einem geeigneten Operator gesucht. Neben der Planung kann zudem die Analogie verwendet werden.

Hierbei wird das aktuelle Problem mit in der Vergangenheit gelösten Problemen verglichen und dementsprechend Strategien zur Problemlösung aktiviert (Klahr, 2000).

Abschließend soll festgehalten werden, dass Erkenntnisgewinnung als Problemlösen die abstrakten Vorstellungen der Naturwissenschaften und das Wissen über die Naturwissenschaften einschließt. Des Weiteren ist Erkenntnisgewinnung aus dieser Perspektive betrachtet stark von deklarativem Wissen und darauf beruhenden Entscheidungen für Prozeduren beeinflusst. Dennoch weist diese Herangehensweise einige Analogien mit Erkenntnisgewinnung als biologiespezifisches prozedurales Wissen auf. Daher wird im Folgenden eine Synthese dieser beiden vorgenommen.

2.1.4.3 Synthese von Erkenntnisgewinnung als biologiespezifisches prozedurales Wissen und als biologiespezifisches Problemlösen

Prozedurales Wissen beinhaltet weder nur die automatisierten Prozesse noch solche, die in letzter Konsequenz zu diesen generiert werden müssen. Vielmehr sind im Kontext von Prozesswissen in der Biologie die Sekundärprozesse vorherrschend. Es könnte sich in diesem Rahmen also um Prozesse handeln, die noch stark mit dem deklarativen Wissen verknüpft sind. Diese Sekundärprozesse haben große Ähnlichkeit mit den beschriebenen Problemlöseprozessen, in denen deklaratives Wissen benötigt wird, um mittels geeigneter Prozeduren das gegebene Problem zu lösen. So kann ein Experiment sowohl als eine Problemlösestrategie angesehen werden als auch als ein Schritt-für-Schritt Prozess in der Konzeptualisierung des prozeduralen Wissens der Kognitionspsychologie. Diese Zusammenführung ist konsistent mit der Annahme von Anderson, dass Problemlöseprozesse der Ursprung prozeduralen Wissens sind (s. Seite 55). Als weiteres Indiz für die Zusammenführung der beiden Ansichten kann der moderierende Effekt des Problemlösens auf den iterativen Erwerb von prozeduralem und konzeptuellem Wissen in der Mathematik gesehen werden (Schneider et al., 2011, s. Seite 59).

Die folgende Abbildung 7 veranschaulicht die Struktur der Wissensarten in der Biologie, die für das Lösen naturwissenschaftlicher Probleme benötigt werden.

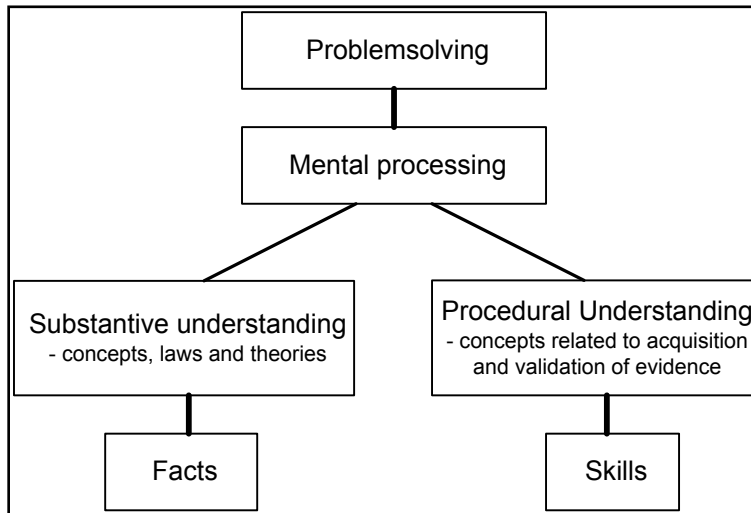


Abbildung 7: Ein Modell für die Naturwissenschaften von Gott und Duggan (1995)

Aus „Procedural understanding in biology: the thinking beyond the doing, von R. Roberts, 2001, *Journal of Biological Education*, 53, S. 113. Copyright 2001 bei Rosalyn Roberts. Wiedergabe mit Genehmigung.

Auch anhand dieser Abbildung wird ersichtlich, dass zwei zu unterscheidende Wissensarten angenommen werden: das *substantive understanding*, das analog zum Teilbereich Konzeptwissen von Kompetenz in Biologie gesehen werden kann, und das *procedural understanding*, welches mit dem Teilbereich Prozesswissen gleichzusetzen ist. Diese spielen in der aktuellen naturwissenschaftlichen oder biologischen Problemsituation zusammen (Shamos, 2002), um zu einem erwünschten Endzustand zu gelangen. Diese Strukturierung steht im Einklang mit dem Befund, dass Prozesswissen (in der Mathematik) die Handlungssequenzen für das Problemlösen bietet (Rittle-Johnson & Wagner Alibali, 1999). Bezogen auf die biologie- bzw. naturwissenschaftsspezifische Erkenntnisgewinnung wurden unterschiedliche Aspekte proklamiert. So wurde der Prozess der Erkenntnisgewinnung in die drei Fähigkeiten (a) Hypothesen formulieren, (b) Variablen identifizieren sowie (c) Daten interpretieren (Klahr, 2000) oder in die zwei Aspekte (a) Planung eines Experiments und (c) Daten sammeln und analysieren unterteilt (Gott & Duggan, 2001, zitiert nach Möller et al., 2009). In einer deutschen Untersuchung im Rahmen des BMBF-Projekts Biologie im Kontext (BiK) an einer Stichprobe von 1 129 Schülerinnen und Schülern der Sekundarstufe I konnten vier zentrale Prozeduren ermittelt werden (Mayer et al., 2008). Hierzu bearbeiteten Schülerinnen und Schüler innerhalb eines Multi-Matrix Designs (s. Seite 103) 24 offene Testaufgaben, die die vier Teilkompetenzen gleichmäßig abdeckten. Der Modellfit eines spezifizierten vierdimensionalen Modells zeigte Evidenz für das Vorliegen der Dimensionen (a) naturwissenschaftliche Fragen formulieren, (b) Hypothesen generieren, (c) Untersuchungen planen sowie (d) Daten analysieren und Schlussfolgerungen ziehen. Die Dimensionen korrelieren untereinander zwischen $r = .33$ und $r = .73$.

(Mayer et al., 2008). Des Weiteren konnte in der Biologie gezeigt werden, dass zwei Aspekte des Experimentierens untersucht werden können: der Aspekt Suche im Hypothesenraum und Analyse von Evidenzen einerseits sowie der Aspekt Testen von Hypothesen andererseits (Hammann et al., 2007).

Die Ausführungen haben gezeigt, dass Erkenntnisgewinnung aus unterschiedlichen Perspektiven betrachtet werden kann. Zum einen beinhaltet Erkenntnisgewinnung Bezüge zum Problemlösen, zum anderen Bezüge zu den Sekundärprozessen des Prozesswissens. Diese beiden Ansichten können integriert werden, da in der Kognitionspsychologie Problemlöseprozesse auch zu den Sekundärprozessen des prozeduralen Wissens gezählt werden können. Um naturwissenschaftliche Probleme zu lösen, also naturwissenschaftliche Methoden der Erkenntnis anzuwenden, müssen Individuen das Konzept- und das Prozesswissen aktivieren. Inwiefern diese beiden spezifisch auf die Biologie bezogenen Wissensarten empirisch voneinander trennbar sind, kann noch nicht bestimmt werden und wird in der vorliegenden Arbeit untersucht. Zunächst werden jedoch Kompetenzen in den Naturwissenschaften und in der Biologie aus der dritten – der differentialpsychologischen Perspektive – betrachtet.

2.2 Die Naturwissenschaften aus differentialpsychologischer Perspektive

Es besteht die Notwendigkeit, Testinstrumente vor ihrem Einsatz für inhaltliche Untersuchungen einer Validierung zu unterziehen. Daher wird die Durchführung von arbeitsteiligen Untersuchungen vorgeschlagen (Schneider & Stern, 2009). Während einige Studien ausschließlich zur Validierung von Messinstrumenten durchzuführen sind und Standards für die eigentliche Testung der jeweilig zu untersuchenden Kompetenz oder des Konstrukts setzen, sollten andere Studien diese validen und reliablen Messinstrumente anwenden. Diese Unterscheidung von Untersuchungen zu Grunde legend, kann die vorliegende Arbeit als eine Validierungsstudie aufgefasst werden. Somit werden neben dem Hauptfokus der Arbeit – der Unterscheidung der Teilbereiche Konzept- und Prozesswissen von Kompetenz in Biologie – weitere Aspekte zur Validierung hinzugezogen. Für eine Vielzahl der in Deutschland neu entwickelten Kompetenztests im Zusammenhang mit dem Paradigmenwechsel zur Outputorientierung liegen nur wenige Studien bezogen auf deren Validierung vor (z. B. Winkelmann, 2009 [für Mathematik in der Grundschule]; Leucht, 2010 [für Englisch als Fremdsprache in der Sekundarstufe I]). Da für die Naturwissenschaften und insbesondere für die Biologie kaum Untersuchungen zu den Teilbereichen Konzept- und Prozesswissen vorliegen, müssen neben der Untersuchung der internen Struktur auch Aspekte konvergenter und diskriminanter Validität aufgenommen werden. Hierdurch kann zum einen die interne Struktur eingehender untersucht werden. Zum anderen wird durch den Einblick in die Zusammenhänge des Tests mit externen

Konstrukten eine erste Einschätzung möglich, welche Fähigkeiten die Kompetenzteilbereiche in Biologie determinieren. Erst darauf aufbauend können Untersuchungen zu diesen beiden Kompetenzaspekten in Biologie folgen. Daher wird im Folgenden die differentialpsychologische Perspektive eingenommen.

In der differentiellen Psychologie werden die Ursachen für stabile und konsistente Unterschiede zwischen Individuen beschrieben und analysiert (Stemmler, Hagemann, Amelang & Bartussek, 2011). Grundlage ist die spezifische Kombination von Persönlichkeitsmerkmalen, die zur Variabilität im Verhalten unterschiedlicher Personen beiträgt. Im Fokus stehen dabei die biologischen, psychologischen und sozialen Entstehungsbedingungen, ihr Einfluss auf das aktuelle Erleben und Verhalten sowie ihre langfristigen physiologischen, psychologischen und sozialen Konsequenzen (Weber & Ramseyer, 2012). Es werden nicht nur Unterschiede zwischen sondern auch innerhalb von Personen untersucht (Stemmler et al., 2011). Für die Validierung von Kompetenz in Biologie ist vor allem der Aspekt des aktuellen Erlebens und Verhaltens – nämlich die Bearbeitung von Testaufgaben zu dieser Kompetenz – von Belang. Wenn Kompetenz in Biologie als eines (von vielen) Persönlichkeitsmerkmalen angesehen werden kann, wird mit differentialpsychologischen Fragestellungen untersucht, welche weiteren Persönlichkeitsmerkmale zu unterschiedlicher Kompetenz in Biologie führen. Die in der vorliegenden Arbeit aufgenommenen Persönlichkeitsmerkmale und verwandten kognitiven Konstrukte werden in einschlägigen Lehrbüchern zur differentiellen und Persönlichkeitspsychologie als verhaltensrelevante Eigenschaften von Personen angenommen (Maltby, Day & Macaskill, 2011; Salewski & Renner, 2009; Stemmler et al., 2011). In vorhergehenden Untersuchungen konnten neben der kognitiven Grundfähigkeit (Brunner, 2006; Rindermann, 2006) und dem Leseverstehen (OECD, 2005; OECD, 2009) auch das Selbstkonzept (Marsh, 1992) und Schulleistungen (Hülür, Wilhelm & Robitzsch, 2011; Winkelmann, 2009) als Prädiktoren für Performanz in Kompetenztestungen in einer spezifischen akademischen Disziplin herausgearbeitet werden. Die Ergebnisse dieser Untersuchungen werden im Folgenden aufgegriffen und in den Rahmen der Validierung von Kompetenz in Biologie eingebettet.

2.2.1 Zusammenhänge naturwissenschaftlicher Kompetenz mit verwandten kognitiven Konstrukten

Schulleistungen und schulisches Lernen hängen besonders stark mit sogenannten proximalen Faktoren zusammen (Wang, Haertel, Geneva & Walberg, 1993). Bei proximalen Variablen handelt es sich um Variablen, die nah an den Schülerinnen und Schülern und an den Prozessen des schulischen Lernens sind. Hierzu zählen die fünf wichtigsten proximalen Variablen: die psychologischen Variablen metakognitiv und kognitiv, die instruktionale Variable Klassenraummanagement sowie die

soziologische Variable des häuslichen Umfelds (Wang et al., 1993). Im Folgenden wird auf die immanenten psychologischen und schulischen Konstrukte kognitive Grundfähigkeit und sprachliche Kompetenzen eingegangen.

2.2.1.1 Der Zusammenhang mit kognitiver Grundfähigkeit

Die kognitive Grundfähigkeit gilt als einer der stärksten, wenn nicht sogar als stärkster Prädiktor für interindividuelle Unterschiede in Schülerleistungen und Wissenserwerbsprozessen (Gustafsson & Balke, 1993; Weinert & Helmke, 1997). Disziplinspezifisches Wissen wird zudem als integraler Bestandteil von kognitiver Grundfähigkeit gesehen (Rolfhus & Ackerman, 1999). Eine typische mittlere Korrelation zwischen kognitiver Grundfähigkeit und akademischer Leistung liegt bei $r = .50$ (Lau & Roeser, 2002). Es wurden sogar Zweifel geäußert, ob die Messung von Schülerleistungen einen Mehrwert im Vergleich zur Messung kognitiver Grundfähigkeit leistet (Rindermann, 2006). Jene Zweifel zeugen von der Bedeutung der Beachtung von Maßen der kognitiven Grundfähigkeit bei der Validierung von Instrumenten zur Erhebung disziplinspezifischer Schülerleistungen. Da auch die in dieser Arbeit zu Grunde liegende Definition von Kompetenz in Biologie stark mit Definitionen kognitiver Grundfähigkeit überlappt, wird diese mit aufgenommen.

Eine konsensualisierte Definition von kognitiver Grundfähigkeit ist noch nicht elaboriert worden. Über gewisse Aspekte dieser Fähigkeit besteht allerdings Einigkeit. Zum einen besteht unter Kognitionspsychologinnen und -psychologen Konsens darüber, dass kognitive Grundfähigkeit aus einem übergeordneten Generalfaktor g und weiteren miteinander in Beziehung stehenden Gruppenfaktoren besteht (Carroll, 1993; Vernon, 1950). Zum anderen unterschrieben 52 Intelligenzforscherinnen und -forscher einen Definitionsversuch, demnach es sich um „*a very general mental capability*“ handelt, „*that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. Rather, it reflects a broader and deeper capability for comprehending our surroundings – ‘catching on,’ ‘making sense’ of things, or ‘figuring out’ what to do.*“ (Gottfredson, 1997, S. 13). Nach einer weiteren Definition ermöglicht die kognitive Grundfähigkeit die Ausführung von Aktivitäten, die durch Schwierigkeit, Komplexität, Abstraktion, Sparsamkeit, Anpassungsfähigkeit und sozialen Wert gekennzeichnet sind. Es handelt sich demzufolge um eine komplexe, dekontextualisierte Fähigkeit zu denkgestütztem Lösen von kognitiven Aufgaben und Problemen (Baumert, Lüdtke, Trautwein & Brunner, 2009). Die kognitiven Aufgaben und Probleme sind hierbei allein durch Wissensabruf nicht erfolgreich lösbar, da sie neuartiger Natur sind und der Problemlöserin bzw. dem Problemlöser das spezifische Wissen fehlt. Daher müssen Problemlöserinnen und Problemlöser Strukturen, Beziehungen, Sinnzusammenhänge und Bedeutungen herstellen, Abstrahieren und Einsichten gewinnen können, also Fähigkeiten des induktiv und deduktiv schlussfolgernden Denkens aufweisen

(Brunner, 2006; Rindermann, 2006). Cattell (1987) beschrieb kognitive Grundfähigkeit als einen Ausdruck des Komplexitätsniveaus von neuartigen Beziehungen, die ein Individuum wahrnehmen und auf das es reagieren kann. Es handelt sich also um eine kognitive Fähigkeit im Umgang mit neuen Problemen, die starke Anforderungen an das schlussfolgernde Denken und geringe Anforderungen an das fachspezifische Vorwissen stellt (Baumert, Brunner, Lüdtke & Trautwein, 2007).

Intelligenzforscherinnen und -forscher untersuchten und untersuchen insbesondere die Struktur der kognitiven Grundfähigkeit. Spearman (1904, 1927) verfasste bereits Ende des 19. Jahrhunderts seine Theorie der allgemeinen kognitiven Leistungsfähigkeit bzw. Theorie der allgemeinen kognitiven Grundfähigkeit *g*. In diesem Zusammenhang stellte er die Theorie der positiven Mannigfaltigkeit aller intelligenten Leistungen auf. Sie besagt, dass akademische (schulische und intellektuelle) Leistungen zum einen aus einer durch deren Interkorrelationen zu bestimmenden, gemeinsamen Dimension – der generellen kognitiven Grundfähigkeit *g* – bestehen. Zum anderen ist *g* unterteilt in verschiedene für jede einzelne Leistung eigene, spezifische Faktoren. Thorndike, Bregman, Cobb und Woodward (1927) veröffentlichten zeitgleich einen alternativen Strukturierungsvorschlag der kognitiven Grundfähigkeit und beschrieben sie als ein Bündel von unabhängigen Fähigkeiten. Diese wurden 10 Jahre später als sieben primäre Gruppenfaktoren identifiziert. Im Deutschen werden diese Fähigkeiten als (a) Gedächtnis, (b) Rechengewandtheit, (c) Auffassungsgeschwindigkeit, (d) schlussfolgerndes Denken, (e) Raumvorstellung, (f) Wortverständnis und (g) Wortflüssigkeit bezeichnet (Rost, 2009). Diese sieben Faktoren sind mit unterschiedlichem Anteil an einer intellektuellen Leistung beteiligt (Thurstone, 1938). Heutzutage wird von einer Abhängigkeit der Fähigkeiten ausgegangen. Jeder Primärfaktor besteht aus einer Kombination eines unabhängigen Primärfaktors und einem zusätzlichen Generalfaktor der intellektuellen Leistungsfähigkeit (Thurstone & Thurstone, 1941). Im weiteren Verlauf wurden weitere hierarchische Modelle der kognitiven Grundfähigkeit veröffentlicht (Burt, 1949; Vernon 1950). In dem Modell von Vernon (1950) wurden die bestehenden Vorstellungen (Spearman, 1927; Thurstone & Thurstone, 1941) vereint; die Mehrzahl der modernen Theorien zur kognitiven Grundfähigkeit nimmt Bezug auf dieses Modell (Rost, 2009). Zusätzlich zu einem Generalfaktor bilden mehrere Gruppenfaktoren die kognitiven Grundfähigkeiten von Menschen ab. Die unzähligen Faktoren auf unteren Ebenen vereinen sich auf jeweils höheren Ebenen zu Gruppen. Auf einer hohen Hierarchiestufe werden die zwei breiten Gruppenfaktoren verbal/numerisch/schulisch – v:ed – und praktisch/mechanisch/räumlich/physikalisch – k:m – angenommen (Vernon, 1950).

Cattell (1941, 1943) vertrat die Hypothese, dass kognitive Grundfähigkeit aus den zwei breit gefassten Facetten fluide und kristalline kognitive Grundfähigkeit besteht, welche stark miteinander in Verbindung stehen. Fluide kognitive Grundfähigkeit umfasst dabei die prozess- und

operationsbezogenen Fähigkeiten. Sie steht in starkem Zusammenhang zur Fähigkeit, sich an neue Situationen anzupassen, daher wird ihr auch eine größere Wichtigkeit beim Lernen auf thematisch neuen Gebieten zugesprochen (Rost, 2009). Die kristalline kognitive Grundfähigkeit besteht aus akkumulierten Lernergebnissen in bestimmten Themenfeldern und korreliert höher mit allgemeinen Fähigkeiten, die durch wiederholte Lernerfahrungen entstanden sind.

Die am weitesten verbreitete und akzeptierte Drei-Schichten-Theorie von Carroll (1993, 2005) unterscheidet drei hierarchisch getrennte Strata. Der ersten Schicht mit 70 Primärfaktoren folgt eine zweite Schicht mit den acht General-Sekundärfaktoren. Dies sind (a) fluide und (b) kristalline kognitive Grundfähigkeit, (c) allgemeine Gedächtnisfähigkeit, (d) breite visuelle Wahrnehmung, (e) breite auditive Wahrnehmung, (f) breite Abruffähigkeit, (g) breite kognitive Schnelligkeit und (h) Verarbeitungsgeschwindigkeit. Die dritte Ebene bildet das schon bekannte *g*.

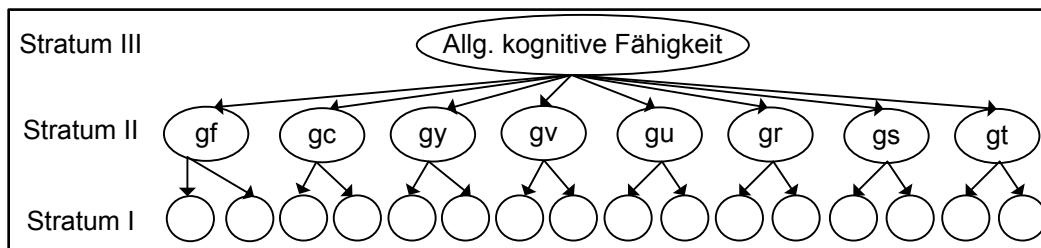


Abbildung 8: Die Drei-Stratum-Theorie von Carroll

Anmerkungen. gf = fluide Intelligenz, gc = kristalline Intelligenz, gy = allgemeine Gedächtnisfähigkeit, gv = breite visuelle Wahrnehmung, gu = breite auditive Wahrnehmung, gr = breite Abruffähigkeit, gs = breite kognitive Schnelligkeit, gt = Verarbeitungsgeschwindigkeit.

Die 70 Primärfaktoren des Stratum I wurden der Übersichtlichkeit nicht vollständig eingezeichnet.

Die kognitive Grundfähigkeit ist für den Wissenserwerb innerhalb einer spezifischen Disziplin von herausgehobener Bedeutung (Weinert, 1992; Weinert & Helmke, 1997). So werden zwischen kognitiver Grundfähigkeit, Prozess-, Konzept- und Strategiewissen Überlappungen angenommen (Rindermann, 2006). Es wird sogar angenommen, dass Wissen in einer bestimmten Disziplin ein integraler Bestandteil der kognitiven Grundfähigkeit sein könnte (Rolfhus & Ackerman, 1999). Die sogenannte Rindermann-Debatte stellt für den deutschen Raum den aktuellsten Diskurs zum Zusammenhang von Maßen der kognitiven Grundfähigkeit und Schülerleistungen dar. Die Debatte zeigt zudem, warum neu entwickelte Schulleistungs- und Kompetenztests mit Maßen der kognitiven Grundfähigkeit validiert werden sollten. Sie wird daher folgend kurz skizziert. Rindermann (2006) veröffentlichte in seiner Kritik an den Messinstrumenten der PISA-Studie und internationaler Schulleistungsuntersuchungen seine Überlegungen zu den starken Zusammenhängen zwischen

Maßen der kognitiven Grundfähigkeit und unterschiedlichen Fähigkeiten und Kompetenzen. Diese können nach seiner Auffassung in einem globalen *g*-Faktor abgebildet werden. Mit Schulleistungsstudien, die den vom deklarativen Wissen abzugrenzenden Kompetenzansatz vorangetrieben haben, werde somit eine einheitliche und bildungsunabhängige Fähigkeit erfasst. Die Re-Analyse des PISA-2000 Datensatzes zeigte starke Zusammenhänge zwischen naturwissenschaftlicher Kompetenz und kognitiver Grundfähigkeit ($r = .90$). Rindermanns Ergebnisse wurden von den PISA-Beauftragten für Deutschland kommentiert. Sie kommen zu dem Schluss, dass Rindermanns These nicht haltbar ist (Baumert et al., 2007; Prenzel, Walter & Frey, 2007). Die starken Korrelationen und Zusammenhänge werden nicht angezweifelt, da kognitive Grundfähigkeit als eine bedeutende Determinante des Wissenserwerbs angesehen wird (Baumert et al., 2007). Weitere Folgestudien konnten diese Argumentation untermauern (Brunner, 2006; Schroeders, Buchholtz, Formazin & Wilhelm, in Druck; Winkelmann, 2009). Daher wird kognitive Grundfähigkeit in Validierungsstudien vielmehr im Sinne einer konvergenten denn einer divergenten Validität aufgenommen.

Personen, die hohe allgemeine kognitive Fähigkeiten bzw. hohe kognitive Fähigkeiten in einer akademischen Disziplin vorzuweisen haben, werden tendenziell auch in anderen Disziplinen eine bessere Performanz zeigen. So konnte in einer für England repräsentativen Untersuchung mit 80 000 Schülerinnen und Schülern für 25 unterschiedliche akademische Disziplinen gezeigt werden, dass kognitive Grundfähigkeit über alle akademischen Disziplinen hinweg eine wichtige Rolle spielt (Deary Strand, Smith & Fernandes, 2007). Die kognitive Grundfähigkeit mit elf Jahren korrelierte mit Ergebnissen in nationalen Examina im Alter von 16 Jahren zu $r = .81$. In der Biologie liegt dieser Zusammenhang bei $r = .51$ (in Physik bei $r = .50$ und in Chemie bei $r = .46$). Auch in Deutschland wurde anhand einer Stichprobe von 137 Berliner Gymnasiastinnen und Gymnasiasten eine Korrelation von $r = .31$ zwischen der Note in den Naturwissenschaften und kognitiver Grundfähigkeit berichtet (Süß, 2007). Eine weitere Untersuchung zu verschiedenen Darbietungsmodi bei der Bearbeitung von naturwissenschaftlichen Inhalten konnte zeigen, dass die fluide kognitive Grundfähigkeit 64 % der Varianz im Verständnis von naturwissenschaftlichen Texten (und Videos) aufklären konnte (Schroeders et al., in Druck). Im Gegensatz dazu klärt das naturwissenschaftliche Wissen der 216 Schülerinnen und Schüler 28 % dieser Varianz auf. Der Zusammenhang zwischen naturwissenschaftlicher Kompetenz und kognitiver Grundfähigkeit fällt dabei stärker aus als der Zusammenhang zwischen Noten in den naturwissenschaftlichen Fächern und kognitiver Grundfähigkeit. Letzterer beträgt $r = .24$ gegenüber $r = .67$ für die naturwissenschaftliche Kompetenz (Lau & Roeser, 2002).

Es wird deutlich, wie stark allgemeine kognitive Fähigkeiten mit den disziplinspezifischen Fähigkeiten zusammenhängen und dass sie folglich in Validierungsstudien Beachtung finden müssen. In der vorliegenden Arbeit wird Kompetenz in Biologie in die beiden Teilbereiche Konzept- und Prozesswissen unterteilt. Die theoretische Herleitung der beiden Teilbereiche verdeutlicht, dass für diese beiden Kompetenzdimensionen unterschiedliche Zusammenhänge angenommen werden könnten. Erkenntnisgewinnung in der Biologie wurde sowohl mit prozeduralem Wissen als auch mit Problemlösen in Verbindung gebracht (s. Kapitel 2.1.4). Dieser Teilbereich könnte daher zu einem höheren Anteil kognitive Ressourcen der Schülerinnen und Schüler beanspruchen als das Konzeptwissen. Testungen der kognitiven Fähigkeit erfordern wenig deklaratives Wissen (Bond 1989). Für die schriftlich erfasste, analytische und allgemeine Problemlösefähigkeit konnte gezeigt werden, dass sie eng mit der kognitiven Grundfähigkeit zusammenhängt (Klieme et al., 2001). Im PISA-Feldtest korrelierten die Maße zum Problemlösen im Median zu $r = .58$ mit der kognitiven Grundfähigkeit (Klieme et al., 2001). Das schlussfolgernde Denken liegt sowohl der allgemeinen als auch der disziplinbezogenen Problemlösekompetenz zu Grunde. Ein hoher Zusammenhang könnte damit erklärt werden, dass insbesondere das Problemlösen – ob allgemein oder disziplinspezifisch konzeptualisiert – kognitive Anforderungen an das Individuum stellt. Bezogen auf diese Zusammenhänge besteht innerhalb der fachdidaktischen Forschung ein Mangel.

Für das Fach Biologie liegen zum Zusammenhang kognitiver Grundfähigkeit mit unterschiedlichen Wissensarten keine Untersuchungen vor. Der Zusammenhang zwischen kognitiver Grundfähigkeit und der Qualität der Lösungen komplexer Probleme in der Biologie ist noch nicht nachgewiesen (Mayer, 2007). Bezogen auf das Prozesswissen kann nur die bereits erwähnte Studie für die Chemie herangezogen werden (Klos et al., 2008). Der Zusammenhang kognitiver Grundfähigkeit – erhoben mit dem kognitiven Fähigkeitstest (KFT) – mit den naturwissenschaftlichen Arbeitsweisen ist als moderat zu bezeichnen (Korrelation in der siebten Klasse $r = .36$, Korrelation in der 12. Klasse $r = .46$). Er bleibt also im mittleren Bereich. Da zu den Zusammenhängen naturwissenschaftlicher Kompetenz mit kognitiver Grundfähigkeit nur wenige Ergebnisse vorliegen, wird erneut auf die benachbarte Disziplin Mathematik zurückgegriffen. Für den deutschen Raum liegen detaillierte empirische Untersuchungen vor (Brunner, 2006; Winkelmann, 2009).

Exkurs kognitive Grundfähigkeit und Leistungen in der Mathematik

In der Dissertation von Brunner (2006) wird unter anderem der Zusammenhang zwischen Maßen der kognitiven Grundfähigkeit und mathematischer Kompetenz anhand der Daten der PISA-Studie 2000 näher untersucht (29 386 Neuntklässler aus 1 301 Schulen). Die theoretischen Vorannahmen, dass mathematische Problemlöseprozesse stark mit allgemeiner und fluider kognitiver Grundfähigkeit überlappen und dass zur Lösung eines Problems (a) Organisieren bzw. Strukturieren, (b) Generalisieren, (c) Adaptieren und (d) Anwendung von Wissen in neuen Kontexten – ergo kognitive Fähigkeiten – benötigt werden, führten zur Spezifikation entsprechender Strukturgleichungsmodelle (Brunner, 2006). Eine theoretische Annahme ist: je stärker die allgemeine kognitive Fähigkeit ausgeprägt ist, desto einfacher fällt einer Person das Einlassen auf neue (komplexe) mathematische Probleme, das Entwickeln effektiver Problemlösestrategien und das Erkennen lösungsrelevanter Regeln. Von den von Brunner (2006) spezifizierten genesteten Strukturgleichungsmodellen zeigte jenes Modell den besten Modellfit, das interindividuelle Unterschiede in den mathematischen Schülerleistungen durch die allgemeine kognitive Fähigkeit, eine mathematikspezifische Fähigkeit und stoffgebietsabhängige Fähigkeiten erklärt (Modell 7a, ebd., S. 94-95). Dieses Modell erklärt im Vergleich zu den weiteren spezifizierten Modellen die interindividuellen Unterschiede in mathematischen Schülerleistungen durch das Zusammenwirken der allgemeinen und disziplinbezogenen Fähigkeiten; die verbale Fähigkeit stellte keinen bedeutenden Erklärungsfaktor dar. Allerdings konnte in einer weiteren Untersuchung anhand der Daten aus (a) der Erhebung zum Lesen- und Mathematikverständnis – Entwicklungen in den Jahrgangsstufen vier und sechs in Berlin (ELEMENT), (b) der Studie Bildungsprozesse und psychosoziale Entwicklung im Jugend- und frühen Erwachsenenalter (BIJU) und (c) der PISA-Studie gezeigt werden, dass bei Kontrolle von Vorwissen die kognitive Grundfähigkeit am Ende der Grundschule nur noch mit 6 % (Viert- bis Fünftklässler) bzw. 8 % (Neunt- bis Zehntklässler) zur Erklärung der Varianzen in den Mathematikergebnissen beiträgt (Baumert et al., 2009). Weitere Ergebnisse liegen im Grundschulbereich vor. In Strukturgleichungsmodellen, in denen Intelligenzmaße und Maße der fachspezifischen mathematischen Fähigkeiten modelliert wurden, hingen diese beiden latenten Faktoren mit $r = .78$ zusammen (Winkelmann, 2009).

Die kognitive Grundfähigkeit ist also eine wichtige übergreifende Kompetenz für fachspezifische Fähigkeiten in Mathematik und den Naturwissenschaften. Neben der kognitiven Grundfähigkeit spielten – wie bereits erwähnt – die sprachlichen Kompetenzen eine weitere wichtige Rolle beim Wissens- und Kompetenzerwerb sowie bei der Testung dieser erworbenen Kompetenzen. So konnte gezeigt werden, dass sprachliche Kompetenzen besser zur Vorhersage geeignet sein könnten als das

abstrakte schlussfolgernde Denken (Ackerman, Beier & Bowen, 2010). Ergebnisse zu den Zusammenhängen fachspezifischer Kompetenzen und der Lesefähigkeit werden im Folgenden zusammengefasst dargestellt.

2.2.1.2 Der Zusammenhang mit sprachlichen Kompetenzen

Naturwissenschaftliches Wissen und naturwissenschaftliche Kompetenz werden in einem beträchtlichen Umfang über schriftliche Dokumente sowie geschriebene und gesprochene Diskurse erlernt, weitergetragen und dokumentiert (Brown, Reveles & Kelly, 2005; Eschenhagen et al., 1998). Insbesondere außerhalb der Schule begegnen Personen den Naturwissenschaften eher über Lesen, Schreiben, Hören und Sprechen als über eigene naturwissenschaftliche Untersuchungen. Zudem spielt in den Naturwissenschaften die Theorie eine wichtige Rolle; naturwissenschaftliche Ideen sind kulturelle Errungenschaften, die Schülerinnen und Schüler nur über Sprache vermittelt werden können (Ford, Brickhouse, Lottero-Perdue & Kittleson, 2006). Dies bedeutet, dass naturwissenschaftliche Kompetenz nicht ohne ein gewisses Maß an akademischen und alltagssprachlichen Kompetenzen erworben werden kann (Fischer et al., 2003; Rost et al., 2004; Taboada & Rutherford, 2011; Wellington & Osborne, 2001).

Cummins (2008) unterscheidet zwischen *basic interpersonal communicative skills* (BICS) und *cognitive academic language proficiency* (CALP). CALP – also akademische Sprache – entsteht vom Zeitpunkt der Geburt an und wird in den frühen Beschulungsjahren aus vorhandenem BICS differenziert (Cummins, 2008). Sie wird daher auch als die Sprache der Beschulung dargestellt und weist Unterschiede zu anderen Sprachen bezüglich des Vokabulars, der Syntax und der Diskurseigenschaften auf (Taboada & Rutherford, 2011). Alltagssprache ist jedoch nicht unabhängig von Fachsprachen, vielmehr ist sie eine Voraussetzung für letztere (Eschenhagen et al., 1998).

Die naturwissenschaftliche Sprache ist eine akademische Sprache, die andere Merkmale als narrative Texte, die Alltagssprache – also BICS – oder die soziale Sprache in der Schule aufweist (Cummins, 2008; Taboada & Rutherford, 2011). So müssen naturwissenschaftlich literate Personen die Fähigkeit besitzen, naturwissenschaftliches Wissen linguistisch zu enkodieren und die spezifisch naturwissenschaftliche, stärker formalisierte Sprache reflektieren können (Westby & Torres-Velásquez, 2000). Die Sprache der Naturwissenschaften ist im Besonderen durch die erschwerte Möglichkeit der Visualisierung oder bildlichen Darstellung gekennzeichnet (Westby & Torres-Velásquez, 2000). Die wichtige Rolle der Sprach- und Lesefähigkeiten zeigt sich auch für die Kommunikations- und Dokumentationsprozesse beim Erlangen von Prozesswissen (Westby & Torres-Velásquez, 2000). Zudem fordert jede Kompetenzmessung in Form eines *paper-and-pencil* Tests die Lese- und Schreibfähigkeit der Testperson (Kuglemeyer & Schecker, 2007). Daher müssen allgemeine

sprachliche Kompetenzen als disziplinübergreifende Fähigkeit in eine Validierungsstudie zu disziplinspezifischer Kompetenz Eingang finden.

Die schriftliche Form der Testung ist also nicht als realitätsfern anzusehen. Auch im alltäglichen Leben kann eine Vielzahl an Problemen nicht gelöst werden, ohne Informationen in Text und Schrift entschlüsseln zu müssen. So ist auch schulischer Unterricht in allen Fächern eine sprachliche Veranstaltung und sprachliche Kompetenzen eine disziplinübergreifende Schlüsselkompetenz (Baumert, 2002). Dennoch ist es möglich, dass bestimmte Aufgabentypen (z. B. erklärenden Charakters) in besonderem Maße mit verbalen Fähigkeiten zusammenhängen könnten (Schneider & Stern, 2009; Schneider et al., 2011). Eine eingehende Untersuchung der sprachlichen Kompetenzen scheint also im Zusammenhang mit der Testung von Kompetenz unabdingbar.

Da in den PISA-Studien sowohl die naturwissenschaftliche als auch die Lesekompetenz untersucht wurde, ist ein für Deutschland repräsentativer Überblick über die Zusammenhänge dieser beiden Kompetenzen möglich. In Tabelle 13 werden für die Erhebungswellen 2000 bis 2009 die Korrelationen mit dem Lesetest für die OECD-Länder und für alle teilnehmenden Länder berichtet.

Tabelle 13: Messfehlerbereinigte Korrelationen der Naturwissenschaftstests mit den Lesetests der PISA Untersuchungen 2000, 2003, 2006 und 2009

Lesetest	Naturwissenschaften			
	PISA 2000	PISA 2003	PISA 2006	PISA 2009
OECD Länder	.80	.83	.84	.87
Teilnehmerstaaten	-	.82	.83	.87

Quellen: Adams & Wu, 2002; OECD, 2005; 2009, 2012

Anmerkungen. - = keine Angaben, PISA = *Programme for International Student Assessment*, OECD = *Organisation for Economic Co-operation and Development*.

Es wird deutlich, dass das Leseverstehen einen starken Zusammenhang mit fachspezifischer Kompetenz aufweist. Die latenten Korrelationen sind über alle Untersuchungen hinweg zwischen $r = .80$ und $r = .87$. Dieser Zusammenhang ist zuvörderst mit dem Umstand zu erklären, dass Aufgaben im *paper-and-pencil* Format von den Schülerinnen und Schülern dekodiert, inhaltlich verstanden und abstrahiert werden müssen. Zudem werden beide Kompetenzen auf einem sehr breiten Spektrum getestet. Auch die Problemlöseindikatoren der PISA Untersuchung, welche thematisch nah am Prozesswissen liegen, wurden im PISA-Feldtest mit der Lesekompetenz korreliert (Klieme et al., 2001). Die mittlere Korrelation von $r = .61$ deutete darauf hin, dass das Prozesswissen in den Naturwissenschaften einen geringeren aber immer noch substantiellen Zusammenhang mit der Lesekompetenz aufweisen könnte. Dieser Zusammenhang wurde auch in der Studie von Klos

et. al. (2008) untersucht, wobei die Lesefähigkeit mit dem verbalen KFT erhoben wurde, der vorrangig schlussfolgerndes Denken misst. Es konnte ein moderater Zusammenhang von $r = .26$ mit den naturwissenschaftlichen Arbeitsweisen gezeigt werden. Auch dass Schülerinnen und Schüler mit Migrationshintergrund größere Schwierigkeiten als Schülerinnen und Schüler ohne Migrationshintergrund im Erlernen von sprachlichen, mathematischen und naturwissenschaftlichen Inhalten haben (Westby & Torres-Velásquez, 2000), deutet auf die Wichtigkeit der sprachlichen Kompetenzen in den naturwissenschaftlichen Fächern hin.

Zusammenfassend kann also resümiert werden, dass der Forschungsstand zum Zusammenhang von kognitiver Grundfähigkeit, und sprachlichen Kompetenzen mit disziplinspezifischer Kompetenz starke Zusammenhänge für einen Kompetenztest in Biologie vermuten lässt. Um diese auszumachen und die Validität der Messinstrumente gewährleisten zu können, werden sie in die Kompetenzmessung aufgenommen. Für die Naturwissenschaften und deren einzelne Disziplinen Biologie, Physik und Chemie sind diese mannigfaltigen Beziehungen mit disziplinübergreifenden Fähigkeiten noch nicht herausgearbeitet worden. Neben den Zusammenhängen zwischen Konzept- und Prozesswissen können mit der vorliegenden Studie diese Relationen mit disziplinübergreifenden Fähigkeiten für die Biologie geleistet werden.

2.2.2 Zusammenhänge naturwissenschaftlicher Kompetenz mit Fachnoten und Persönlichkeitsmerkmalen

Zur weiteren Ausschärfung der Teilbereiche Konzept- und Prozesswissen von Kompetenz in Biologie und Validierung des vorliegenden Kompetenztests werden im Folgenden weitere schulnahe und psychologische Konstrukte untersucht. Nach der Überprüfung der Zusammenhänge mit disziplinübergreifenden Konstrukten, steht nun die Untersuchung mit Fachnoten und Selbstkonzepten bezogen auf verschiedene Schulfächer im Fokus. Ein Vergleich der differentiellen Korrelationsmuster hilft bei der weiteren Verortung der Kompetenzteilbereiche von Biologie. In Fachnoten spielen neben der fachlichen Leistung zusätzliche Faktoren wie kognitive Grundfähigkeit, motivationale Aspekte und Interesse, Arbeitstechniken oder die Bezugsnormorientierung hinein (Rost, 2009). Insbesondere motivationale und affektive Variablen haben in den psychologischen Theorien zum Lernen an Bedeutung gewonnen, da sie eine Schlüsselrolle bei der Ausdauer und dem Enthusiasmus für das Lernen innehaben (Wang et al., 1993).

2.2.2.1 Der Zusammenhang mit Fachnoten in naturwissenschaftlichen und nicht-naturwissenschaftlichen Disziplinen

Zensuren „stellen eine Bewertung von schulischen Leistungen dar und geben als solche wieder, wie ein Lehrer die Leistung beurteilt“ (Tent, Fingerhut & Langfeldt, 1976, S. 23). Seit der Mitte des 19. Jahrhunderts haben Zensuren eine Bedeutung für den individuellen Bildungsweg von Schülerinnen und Schülern und tragen zur zukünftigen Stellung in der Gesellschaft bei (Ingenkamp, 1981). Die Zensurengebung hat – im Gegensatz zu einem Schulleistungstest – mehrere gesellschaftlich und individuell wichtige Funktionen: Rangierungs- und Auslesefunktion, Orientierungs- und Berichtsfunktion, Berechtigungsfunktion und Anreizfunktion (Tent et al., 1976). Diese unterschiedlichen Funktionen spiegeln wider, dass Zensuren sowohl ein pädagogisches Instrument als auch ein Selektionsinstrument sind (Trautwein & Baeriswyl, 2007). Mit dem Bedeutungszuwachs nahm auch die Kritik an ihnen zu. International besteht eine lange Tradition der wissenschaftlichen Untersuchung der Beurteilung von Schulnoten. In Deutschland herrschte bis zum Paradigmenwechsel zur Ouputorientierung die Meinung vor, dass Zensuren eine objektive Bewertung einer Schülerin bzw. eines Schülers darstellen. Die Kritik im deutschsprachigen Raum beschränkte sich eher „auf allgemeinere pädagogische und psychologische Erwägungen“ (Ingenkamp, 1981, S. 12) als auf empirische Untersuchungen. Nachdem im angelsächsischen Raum, in Frankreich, in den Niederlanden und in Schweden bereits früh begonnen wurde, psychometrische Messungen der Zensurengebung gegenüberzustellen, wurde in Deutschland Anfang der 70er Jahre eine kritische Auseinandersetzung mit der Validität und Objektivität von Noten vorangetrieben (Ingenkamp, 1995, 1981; Tent et al., 1976). Es wurde eine Reihe von Untersuchungen im deutschsprachigen Raum herangezogen, die die Subjektivität von Lehrerurteilen unterstreicht. So konnte gezeigt werden, dass Mädchen – auch bei Kontrolle der kognitiven Grundfähigkeit – von Lehrerinnen aber auch von Lehrern tendenziell besser bewertet werden (Carter, 1952). Es wurden Vermutungen darüber angestellt, welche weiteren Aspekte Lehrerinnen und Lehrer in ihr Urteil über die Schülerinnen und Schüler mit einfließen lassen.

Die genannten Aspekte können auf Seiten der Lehrkraft und auf Seiten der Schülerin bzw. des Schülers verortet werden. Als Attribute auf Seiten der Lehrkraft werden Berufserfahrungen (Lintorf, 2012), Zustand der Ermüdung bzw. Langeweile während der Benotung sowie Reihenfolge der zu benotenden Dokumente und Einschätzungen zu Persönlichkeitsfaktoren vermutet (Hadley, 1954; Ingenkamp, 1981). Auf Seiten der Schülerin bzw. der Schüler könnten hingegen mehrere Aspekte eine Rolle bei der Zensurengebung spielen (Hadley, 1954; Ingenkamp, 1981). Hierzu zählt der kognitive Faktor allgemeine Intelligenz und die tatsächliche Leistung. Unter den vielen sozialen Faktoren können die Lehrer-Schüler-Beziehung oder Beliebtheit bei der Lehrkraft, das Betragen, die Pünktlichkeit und Teilnahme sowie Gehorsam und soziale Stellung genannt werden. Auch der

biologische Faktor Geschlecht spielt bei der Zensurengebung eine ausschlaggebende Rolle. Des Weiteren wirken persönliche Erscheinung, Anstrengung und Haltung, Handschrift, sprachliche Ausdrucksfähigkeit und Sitzplatz im Klassenraum auf die Beurteilung der Lehrkraft. Zu beachten ist weiterhin, dass Lehrkräfte bei der Benotung auch vorangegangene Leistungen mit einbeziehen. Sie müssen im Laufe eines Schuljahres unzählige Einzelbeobachtungen unterschiedlichster Erfassungsmethoden vornehmen (Ingenkamp, 1995; Lintorf, 2012). Diese Faktoren wurden nicht nur aus rein empirischen Untersuchungen gewonnen sondern beruhen teilweise auf den Einschätzungen der Forscherinnen und Forscher (Ingenkamp, 1995).

Weiterhin konnte gezeigt werden, dass Lehrerinnen und Lehrer dazu neigen, bei ihnen beliebte Schülerinnen und Schüler höher zu bewerten als unbeliebte Schülerinnen und Schüler. Die Sympathie der Lehrkraft gegenüber der Schülerin bzw. des Schülers nahm hierbei den gleichwertigen Stellenwert ein wie die tatsächliche Leistung (Hadley, 1954). Es scheint also eine Wechselwirkung zwischen dem Verhalten der Schülerin bzw. des Schülers und den Einstellungen der Lehrkraft zu geben. Lehrerinnen und Lehrer diagnostizieren somit verhaltens- und prozessorientiert (Ingenkamp, 1981). Lehrerinnen und Lehrer gehen wiederum zu 70 % davon aus, dass ihre Zensurengebung gerecht ist und zu 54 % davon, dass ihre Notengebung durch eigene Stimmungsschwankungen beeinflusst sein könnte (Steltmann, 1977; zitiert nach Ingenkamp, 1981, S. 14). Tatsächlich zensieren Lehrkräfte in den sogenannten Kernfächern (Muttersprache, Mathematik und Fremdsprache) strenger als in den Naturwissenschaften und orientieren ihren Bezugsrahmen stark an dem Klassenmaßstab (Ingenkamp, 1995). So konnte gezeigt werden, dass die Korrelation zwischen Leistungen und curricularen Tests innerhalb einer Klasse hoch ausfällt, dieser Zusammenhang über die Klassen hinweg jedoch nennenswert geringer ist (Tent, et al., 1976).

Die Mehrzahl von Studien zu Lehrerurteilen schulischer Leistungen liegt für die (Mutter-)sprachen und Mathematik vor. In älteren Studien wurden hauptsächlich Korrelationen zwischen Schulnoten und Schulleistungstests sowie Untersuchungen zur Faktorstruktur herangezogen, um die Validität der Noten einschätzen zu können. So wurden Korrelationen zwischen Schulnoten und Schulleistungstests von $r = .36$ bis $.66$ ermittelt (Ingenkamp, 1981). Mehrere Studien konnten zeigen, dass Schulnoten weder objektiv noch reliabel oder zuverlässig sind (Ingenkamp, 1995; Lintorf, 2012; Trautwein & Baeriswyl, 2007). Die faktoranalytische Betrachtung von Schulzeugnisnoten in der Sekundarschule offenbarte einen Fremdsprachenfaktor, einen mathematisch-naturwissenschaftlichen Faktor und einen Sachfaktor (Denig & Weis, 1970). Neuere Untersuchungen zeigen ein hierarchisches Modell mit einem Generalfaktor und fünf Faktoren zweiter Ordnung (Sprachenfaktor, Naturwissenschaftenfaktor, sozial- und geisteswissenschaftlicher Faktor sowie ein praktischer und ein räumlich-praktischer Faktor) (Gustafsson & Balke, 1993). Für diese vier Faktoren zweiter Ordnung

konnte gezeigt werden, dass sie sehr hoch mit einem gemeinsamen Generalfaktor zusammenhängen (Gustafsson & Balke, 1993)

Ein differentieller Vergleich der Schulleistungen in den fünf mathematischen Kompetenzbereichen (a) Zahlen und Operationen, (b) Raum und Form, (c) Muster und Strukturen, (d) Größen und Messen sowie (e) Daten, Häufigkeit und Wahrscheinlichkeit mit den Fachnoten in Mathematik und Deutsch an einer Grundschulstichprobe zeigte Korrelationen zwischen $r = .45$ und $r = .72^5$. Der gemittelte Zusammenhang mit der Fachnote in Mathematik ($r_{MW} = .64$) unterschied sich jedoch nicht sehr stark von dem mittleren Zusammenhang mit den Fachnoten in Deutsch ($r_{MW} = .52$) (Winkelmann, 2009). Zwei weitere Studien zum Zusammenhang von Fachnoten mit Ergebnissen in Schulleistungstest zeigten erneut mittlere bis hohe Zusammenhänge (Gustafsson & Balke, 1993; Trautwein & Baeriswyl, 2007). Im Vergleich zu Zusammenhängen über die Fächer hinweg fallen diese geringer aus (Lüdtke, Köller, Artelt, Stanat & Baumert, 2002). Vergleichende Studien, die systematisch die Fachnoten der naturwissenschaftlichen Fächer untereinander sowie der naturwissenschaftlichen Fächer zu den Fachnoten in Deutsch und Mathematik in Beziehung setzen, liegen nicht vor. Anhand der Daten der TIMS-Studie 2007 konnte für die Kompetenz in Sachkunde in der Grundschule ein Zusammenhang von $r = .50$ mit der Schulnote gezeigt werden (Lintorf, 2012). Die PISA-Studie 2006 bietet Ergebnisse für die drei Fächer Biologie, Physik und Chemie. Hier liegen die Korrelationen zwischen Schulnoten und Kompetenz in der gleichen akademischen Disziplin zwischen $r = .34$ und $.36$ (Schütte, Frenzel, Asseburg & Pekrun, 2007). Des Weiteren wurde in einer Querschnittsuntersuchung keine Korrelation zwischen dem Kompetenzanstieg und der Biologieleistung gefunden (Grube & Mayer, 2009).

Indizien für die Referenzgruppeneffekte der Schulnotenvergabe konnten anhand von Mehrebenenanalysen gezeigt werden, in denen das mittlere Leistungsniveau der Klasse negativ auf das Lehrerbeurteilung wirkte (Trautwein & Baeriswyl, 2007). Auch zu längsschnittlichen Zusammenhängen zwischen Schulnoten und Schulleistungstests liegen aktuelle Befunde vor. An einer Stichprobe von 158 Schülerinnen und Schülern an Berliner Gymnasien konnte gezeigt werden, dass sowohl die Fachnoten als auch die Leistungswerte der Schülerinnen in Mathematik und Deutsch von der achten zur zehnten Klassenstufe zunahmen (Hülür et al., 2011). Verbesserungen in den Leistungen führten also auch zu einer Verbesserung der Zensuren. Während die Verbesserungen in den Testleistungen in Deutsch und Mathematik voneinander unabhängig waren, zeigte die Verbesserung der Fachnoten der beiden Fächer einen Zusammenhang ($r = .32$). Die Fachnoten in Mathematik und Deutsch korrelierten in der achten Klassenstufe durchschnittliche zu $r = .43$ und in der 10. Klassenstufe zu $r = .49$. Die Fachnoten in Mathematik hingen wiederum über die

⁵ In allen aufgeführten Untersuchungen zu den Zusammenhängen mit den Fachnoten wurden die Noten in umgekehrter Form kodiert (6 = höchster / bester Wert, 1 = geringster / schlechtester Wert), sodass berichtete positive Korrelationen einen positiven Zusammenhang in der Realität widerspiegeln.

Untersuchungsgruppen und Messzeitpunkte gemittelt mit den Ergebnissen im Schulleistungstest Mathematik bei $r = .51$ zusammen; der gleiche Zusammenhang liegt für Deutsch bei $r = .42$. Die jeweiligen gekreuzten Korrelationen zwischen Fachnoten in Mathematik und Leistungen im Deutschtest bzw. Fachnoten in Deutsch und Leistungen im Mathematiktest lagen nur zwischen $r = .17$ bis $.42$ (Hülür et al., 2011). Die Ausführungen zu den Zusammenhängen von Fachnoten mit Schulleistungstests zeichnen ein ambivalentes Bild. Zum einen wird deutlich, dass mit der Benotung von Schülerinnen und Schülern schulische Leistungen beurteilt werden. Zum anderen konnte gezeigt werden, dass eine hohe Anzahl an weiteren Merkmalen zu dem abschließenden Urteil der Lehrkräfte führt. Befunde zur Varianzaufklärung von Kompetenztest anhand von Schulnoten konnten in den berichteten Untersuchungen nicht gefunden werden.

2.2.2.2 Der Zusammenhang mit Persönlichkeitsmerkmalen – Die naturwissenschaftlichen Selbstkonzepte

Emotionale Faktoren beeinflussen die Lernbereitschaft, die Lernstrategien sowie selbstbezogene kognitive Prozesse und akademische Leistungen von Schülerinnen und Schülern (PISA-Konsortium Deutschland, 2007). Zusammenhänge von akademischen Leistungen mit Persönlichkeitsmerkmalen, die auf das zu erfassende Konstrukt fokussieren, liegen höher als Zusammenhänge mit Persönlichkeitsmerkmalen, die sehr globale oder weitere Konstrukte fokussieren (Lau & Roeser, 2002; Marsh & Hau, 2003; Möller et al., 2007; PISA-Konsortium Deutschland, 2007). Im Zusammenhang mit Performanzleistungen und Kompetenz werden von verschiedenen Autoren als bedeutende Einflussvariablen für Schülerleistungen neben der kognitiven Fähigkeit die Konstrukte Selbstkonzept und Motivation genannt (Beghetto, 2007; Lau & Roeser, 2002; Weinert & Helmke, 1997). Für das Selbstkonzept konnte gezeigt werden, dass es in den unterschiedlichsten psychologischen Disziplinen wie z. B. neben der Bildungspsychologie auch in der Gesundheits- oder Persönlichkeitspsychologie sowohl als gewünschte Ergebnisvariable als auch als Mediatorvariable eine herausgehobene Rolle spielt (Marsh & Craven, 2006). Die Bedeutung des akademischen Selbstkonzepts für akademische Leistungen wird im Folgenden kurz umrissen.

In einem einflussreichen Artikel von White (1959) wird die Wichtigkeit des Einbezugs von motivationalen Faktoren bei der Bestimmung von Kompetenz herausgestellt. Das motivationale Konstrukt Selbstkonzept ist ein zentraler Bestandteil der Theorien zur Lernmotivation (Gniewosz, 2010) und beinhaltet Gefühle und Wissen über die eigenen Fähigkeiten, Fertigkeiten, die eigene Erscheinung und die soziale Akzeptanz (Zanobini & Usai, 2002). Für das Selbstkonzept wurde ein mehrdimensionales, hierarchisches Modell vorgelegt (Shavelson, Hubner & Stanton, 1976), welches im weiteren Verlauf ausdifferenziert und verändert wurde (Marsh, 1990; Marsh, Lüdtke, Trautwein & Morin 2009; Marsh & Shavelson, 1985, s. Abbildung 9). Selbstkonzept wird als ein hierarchisch

aufgebautes Konstrukt beschrieben; in dieser Hierarchie befindet sich auf der obersten Ebene ein übergreifendes Selbstkonzept. Auf dem darunter liegenden Niveau wird in das akademische und nicht-akademische Selbstkonzept differenziert. Auf der nächst tieferen Ebene werden verschiedene kontextbezogene Subkomponenten voneinander getrennt (Marsh et al., 2009; Weinert, 2001). Das nicht-akademische Selbstkonzept unterteilt sich in die Subkomponenten soziales, emotionales und physisches Selbstkonzept. Das akademische Selbstkonzept wird in zwei Faktoren unterteilt: den mathematischen bzw. naturwissenschaftlichen und den verbalen Faktor (Marsh & Craven, 2006). Das akademische Selbstkonzept basiert sowohl auf dem individuellen akademischen Erfolgsniveau als auch auf einem Vergleich zum Mittel anderer Personen einer sozialen Gruppe (Marsh & Hau, 2003) und den Leistungsfeedbacks aus verschiedenen Quellen (Gwienosz, 2010). Es ist eine wichtige Determinante bezogen auf Kompetenz- und Performanzleistungen innerhalb einer bestimmten akademischen Disziplin (Marsh & Craven 2006), da es im Zusammenhang mit dem akademischen Erfolg, den Leistungsrückmeldungen und der langfristigen Aspiration in dieser Disziplin steht (Gniewosz, 2010; Marsh & Hau, 2003). Das akademische Selbstkonzept wird im schulischen Zusammenhang weiter in Schulfächer differenziert (Weinert & Helmke, 1997), da die akademischen Selbsteinschätzungen sehr stark auf den schulischen Erfahrungen beruhen. Die fachspezifischen Anteile des Selbstkonzepts bilden die psychischen Prozesse daher deutlich besser ab als z. B. ein globales Selbstkonzept (Gniewosz, 2010). Diese hierarchisch-differenzierte Struktur der Selbstkonzepte konnte nur unter Einbezug sogenannter Kernfächer bestätigt werden (Marsh, 1990). Die Unterteilung des akademischen Selbstkonzepts in verschiedene disziplinbezogene Subkomponenten stellt – insbesondere in den letzten Jahren – einen hervorgehobenen Fokus in der Selbstkonzeptforschung dar. Es konnte gezeigt werden, dass diese Subkomponenten stark disziplinbezogen differenziert sind (Marsh, Trautwein, Lüdtke, Köller & Baumert, 2006).

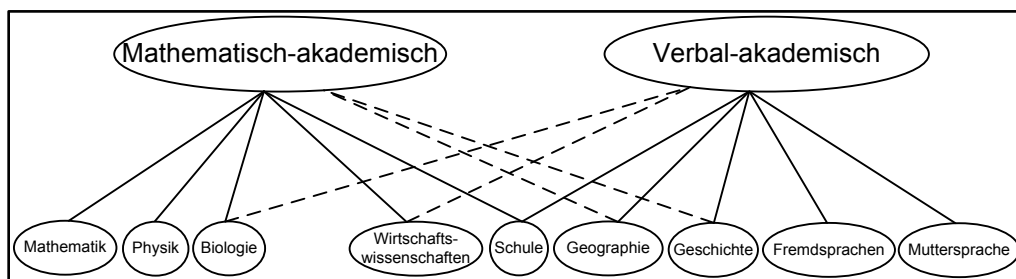


Abbildung 9: Struktur des Selbstkonzepts nach Marsh, Byrne und Shavelson.

Aus "The Structure of Academic Self-Concept: The Marsh/Shavelson Model" von H. W. Marsh, 1990, *Journal of Educational Psychology*, 82, S. 624. Copyright 1990 bei APA und Herbert Marsh.

Wiedergabe mit Genehmigung von Herbert Marsh.

Anmerkungen. Übersetzung v. Verf..

Um Unterschiede in den akademischen Selbstkonzepten verschiedener Disziplinen zu erklären, werden das *internal/external frame of reference* (I/E) Modell und das *reciprocal effects* Modell (REM) herangezogen (Marsh & Köller, 2004). Entsprechend dem I/E-Modell wird das akademische Selbstkonzept einer bestimmten Disziplin im Zusammenspiel von externalen bzw. normativen und internalen bzw. ipsativen Vergleichsprozessen gebildet (Marsh, 1992). Bei dem externalen Prozess vergleichen Schülerinnen und Schüler ihre wahrgenommene eigene Performanz in einer akademischen Disziplin mit der wahrgenommenen Performanz anderer Schülerinnen und Schüler in der gleichen Disziplin. Aufgrund dieses Prozesses müssten die Selbstkonzeptwahrnehmungen unterschiedlicher Disziplinen ungefähr so stark miteinander korrelieren, wie auch die Leistungen in dieser Disziplin miteinander korrelieren. Der interne Prozess zeichnet sich dadurch aus, dass Schülerinnen und Schüler ihre Performanz in einer bestimmten akademischen Disziplin (z. B. im Schulfach Biologie) mit ihrer eigenen Performanz in anderen Disziplinen (z. B. im Schulfach Mathematik) vergleichen. Dieser Vergleichsprozess müsste zu einer negativen Korrelation zwischen Selbstkonzepten unterschiedlicher Disziplinen führen, denn wenn eine akademische Disziplin als eine Stärke empfunden wird, muss eine andere Disziplin konsequenter Weise als eine schwächere empfunden werden. Die Kombination dieser beiden Vergleichsprozesse würde dann zu einer Korrelation der Selbstkonzepte unterschiedlicher Disziplinen von annähernd Null führen (Marsh et al., 2009). Am Beispiel der Selbstkonzepte verbaler Fähigkeiten und mathematischer Fähigkeiten konnte dieses Modell bestärkt werden (Marsh & Craven, 2006; Marsh & Köller, 2004; Marsh et al., 2009). Diese Vergleichsprozesse wurden innerhalb der naturwissenschaftlichen Disziplinen nur wenig untersucht und stehen in keinem Verhältnis zu der Literaturlage bezogen auf verbales und mathematisches Selbstkonzept (Jansen, Schroeders & Lüdtke, in Revision).

In dem Modell der reziproken Effekte – dem sogenannten REM-Modell – wird dargestellt, dass die Konstrukte akademische Leistung und akademisches Selbstkonzept sowohl die Ursache als auch der Effekt für das jeweils andere Konstrukt sind (Marsh & Craven, 2006; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005, 2006). Laut diesem Modell hat die vorangegangene Selbstkonzeptwahrnehmung in einer bestimmten akademischen Disziplin einen positiven Effekt auf die Leistung in dieser Disziplin. Das Reziprok dazu besagt, dass auch die vorangegangene Leistung einen positiven Effekt auf das aktuelle Selbstkonzept in jener Disziplin hat (Marsh, 2002). Dieses Modell ist ein weiteres Indiz für den starken Zusammenhang zwischen disziplinbezogenem Selbstkonzept und Leistungen in dieser Disziplin. Das I/E-Modell und das Modell der reziproken Effekte wurden synthetisiert und die Existenz beider Theorien untermauert (Marsh & Köller, 2004).

Da die kontextbezogenen Selbstkonzepte in einem höheren Zusammenhang mit kontextbezogenen Disziplinen stehen, sollte der Zusammenhang eines Tests zur Erfassung von

Kompetenz in Biologie mit dem Selbstkonzept in Biologie höher ausfallen als die Zusammenhänge mit Selbstkonzepten anderer Disziplinen. Diese Zusammenhänge wurden an Grundschülerinnen und Grundschulern (Weinert & Helmke, 1997), an Fünft- und Sechstklässlern für die Disziplinen Mathematik und Deutsch (Gniewosz, 2010) sowie Zehnt- und Elftklässlern (Lau & Roeser, 2002) untersucht. In der Grundschule lagen die Korrelationen zwischen dem Selbstkonzept und der Testleistung der Schülerinnen und Schüler in Mathematik zwischen $r = .36$ in der zweiten Klasse $r = .50$ und in der vierten Klasse. In der Sekundarschule lagen die Korrelationen zwischen dem Selbstkonzept und den Noten zwischen $r = .34$ und $r = .64$ für das Fach Deutsch und zwischen $r = .45$ und $r = .72$ für das Fach Mathematik. Im Zusammenhang mit der längsschnittlichen Testung des REM-Modells konnte bezogen auf die Fächer Englisch (als Muttersprache) und Mathematik gezeigt werden, dass die vorangegangenen Leistungen in Mathematik positiv mit dem folgenden Selbstkonzept in Mathematik und negativ mit dem folgenden Selbstkonzept in Englisch zusammenhängen (Marsh & Craven, 2006). Der gleiche Effekt konnte in einer weiteren Untersuchung für die Auswirkungen von vorangegangenen Leistungen in Deutsch (als Muttersprache) auf die beiden Selbstkonzepte und für vorangegangene Selbstkonzepte auf die folgenden Leistungen in Mathematik und Deutsch gezeigt werden (Marsh & Köller, 2004). Ähnliche Effekte wurden für die drei naturwissenschaftlichen Fächer noch nicht untersucht. Da Biologie neben dem akademisch-mathematischen auch auf dem akademisch-verbalen Faktor zweiter Ordnung von Selbstkonzept angesiedelt ist und Physik eher auf dem akademisch-mathematischen Faktor lädt, sind auch hier differenzielle Korrelationsmuster der unterschiedlichen Selbstkonzepte zu den naturwissenschaftlichen Disziplinen denkbar (Marsh, 1990). Für diese könnten dann – wie für das mathematische und verbale Selbstkonzept – stabile Muster entstehen,.

Für die Naturwissenschaften liegen nur vereinzelt Ergebnisse zum Zusammenhang zwischen Selbstkonzept und Test- bzw. schulischen Leistungen vor. In der Literatur wird jedoch unterstrichen, dass motivationale Faktoren für die Verbesserung von *scientific literacy* anzunehmen sind (Gross, 2006). Weiterhin bleibt zu erwähnen, dass Zusammenhänge mit den Fachnoten der gleichen akademischen Disziplin systematisch höher ausfallen als Zusammenhangsergebnisse aus Kompetenztestungen (z. B. Jansen et al., in Revision). Des Weiteren zeigte eine Analyse der Zusammenhänge von Wissenstests in 20 akademischen Disziplinen, dass die Biologie im Gegensatz zur Chemie und Physik, neben hohen Ladungen auf einem Naturwissenschaftsfaktor auch hohe Ladungen auf einem geisteswissenschaftlichen Faktor zeigt (Rolfhus & Ackerman, 1999).

In der PISA-Studie 2006 konnte ein Zusammenhang des Selbstkonzept der Schülerinnen und Schüler mit deren Leistungen im Test von $r = .26$ (PISA-Konsortium Deutschland, 2007) gezeigt werden. Bei Schülerinnen und Schülern des 10. und 11. Jahrgangs lag der Zusammenhang mit der

Kompetenz in den Naturwissenschaften bei $r = .54$ (Lau & Roeser, 2002). Pajares, Britner und Valiante (2000) berichten für die Jahresabschlussnote in den Naturwissenschaften von Siebtklässlern eine Korrelation mit dem Selbstkonzept von $r = .43$. Es wurden jedoch auch höhere Zusammenhänge gefunden. So liegt für Siebt- bis Zehntklässler eine Korrelation des Selbstkonzepts mit der Schulnote in den Naturwissenschaften von $r = .70$ vor (Marsh, 1992). In zwei Studien zum Selbstkonzept in den Naturwissenschaften und in der Mathematik wurden sowohl substanzielle Regressionsgewichte in der Vorhersage der naturwissenschaftlichen Testleistung durch das Selbstkonzept in den Naturwissenschaften von $\beta = .34$, in der Vorhersage des Selbstkonzepts durch die Testleistung $r = .54$ (Chiu, 2008, 2012) als auch nur marginale Zusammenhänge für Biologie von $r = .23$ berichtet (Schütte et al., 2007). In einer vergleichenden Untersuchung der Selbstkonzepte in den Naturwissenschaften und Mathematik wurde jedoch die Vermutung aufgestellt, dass die Mathematik als Studium von Mustern und Logik eine wichtige Basis für die Naturwissenschaften sein könnte (Chiu, 2008).

Die umfänglichste Untersuchung zu Selbstkonzept in den Naturwissenschaften wurde von Jansen et al. (in Revision) vorgelegt. Sie konnten die dreidimensionale Struktur der Selbstkonzepte in Biologie, Chemie und Physik aufzeigen. Die Zusammenhänge zwischen diesen naturwissenschaftlichen Selbstkonzepten unterschieden sich; Biologie hing geringer mit Physik zusammen als mit Chemie (s. Tabelle 14).

Tabelle 14: Messfehlerbereinigte Korrelationen zwischen den Selbstkonzepten in Biologie, Chemie und Physik

	Chemie	Physik
Biologie	.43	.34
Chemie	-	.49

Quelle: Jansen et al. (in Revision)

Bezogen auf Zusammenhänge der naturwissenschaftlichen Selbstkonzepte mit akademischen Leistungen konnten sie zeigen, dass für die Fachnoten ein differenzierteres Muster als für die Leistungen in einem Schulleistungstest vorliegt. Die Fachnote in einer akademischen Disziplin konnte einen relativ höheren Anteil des Selbstkonzepts im gleichen Fach erklären. Die Regressionsgewichte lagen zwischen $\beta_{\text{Biologie}} = .33$ und $\beta_{\text{Chemie}} = .42$. Die Regressionsgewichte zu fachfremden Selbstkonzepten tendierten gegen Null. Für Leistungswerte konnte dieses Muster nicht repliziert werden. Das Selbstkonzept Biologie wurde z. B. sowohl durch die Leistungswerte in der Physik ($\beta = .19$) als auch in der Biologie ($\beta = .12$) bestimmt. Insgesamt lagen auch die Varianzaufklärungen durch die Noten mit $R^2_M = .15$ höher als die Varianzaufklärung durch die Leistungswerte ($R^2_M = .08$).

Auch in den Teildisziplinen der Naturwissenschaften ist die Datenlage nicht zufriedenstellend. Bei Untersuchungen des Zusammenhangs von emotionalen Faktoren mit Lernen in Physik wurde ein Zusammenhang von $r = .33$ der Leistung in einem Test zur Elektrizität und Selbstkonzept gefunden (Laukenmann et al., 2000) bzw. ein Zusammenhang von $r = .49$ mit Schulnoten in Physik (Möller, Streblow, Pohlmann & Köller, 2006). In der Biologie wurde eine Korrelation von $r = .49$ zwischen einem Test in Genetik und dem Selbstkonzept berichtet (Thomas et al., 1993). In der Chemie war das Selbstkonzept ein valider Prädiktor für die Fachnoten ($R^2 = .28$) und konzeptuelles Verständnis ($\beta = .50$) (Lewis, Shaw & Heitz, 2009; Nieswandt, 2007).

Mit den theoretischen Ausführungen ist es zum einen gelungen, die interne Struktur von Kompetenz in Biologie näher zu bestimmen. Die Einnahme von Perspektiven aus unterschiedlichen Disziplinen führte zur Differenzierung von Kompetenz in Biologie in das Konzept- und das Prozesswissen. Zum anderen wurde der Forschungsstand zu Zusammenhängen von Kompetenz in Biologie mit relevanten externen Kriterien zusammengefasst und in das Konzept der Validität eingebettet. Der Forschungsstand zeigt, dass für die Naturwissenschaften einige Befunde vorliegen; nur einzelne Untersuchungen fokussieren auf die „Teildisziplin“ Biologie. Mit der vorliegenden Arbeit bietet sich also nicht nur die Möglichkeit, Untersuchungen zu Zusammenhängen von Kompetenz in Biologie mit sprachlichen Kompetenzen oder naturwissenschaftlichen Selbstkonzepten in einen bestehenden Rahmen einzuordnen, vielmehr können neue Wege beschritten werden.

3 Fragestellungen zu Aspekten der Konstruktvalidität von Kompetenz in Biologie

In den beiden vorangehenden Kapiteln wurden in der theoretischen Herleitung des Konstrukts Kompetenz in Biologie sowie durch die Bezugnahme auf disziplinübergreifende Konstrukte Fragestellungen herausgearbeitet, die im Folgenden formuliert werden. Die Darlegung der Fragestellungen ist in zwei große, an die Struktur des theoretischen Teils angelehnte Abschnitte aufgeteilt. Im folgenden Abschnitt wird die fachdidaktisch-kognitionspsychologische Fragestellung entwickelt. Auf dieser Grundlage werden die differentialpsychologischen Fragestellungen aufgegriffen.

3.1 Fragestellung zum Zusammenhang des Konzept- und Prozesswissens in Biologie – Fachdidaktisch-kognitionspsychologische Perspektive

Zunächst wurde die interne Struktur der Kompetenz in Biologie betrachtet. Die Testung der internen Struktur von Kompetenz in Biologie muss in der vorliegenden Arbeit auf die Dimensionen Konzept und Prozesswissen beschränkt bleiben. Wünschenswert wäre neben der Testung auf diese Zweidimensionalität auch eine jeweilige Spezifikation der internen Struktur der Konzept- und Prozessdimension von Kompetenz in der Biologie auf Grundlage der Kompetenzdefinition. Das Konzeptwissen könnte dann bezogen auf die Subkompetenzen (a) System, (b) Struktur und Funktion sowie (c) Entwicklung näher untersucht werden. Das Prozesswissen könnte differenziert nach den Subkompetenzen naturwissenschaftliche Untersuchungen, Modellbildung und wissenschaftstheoretische Reflexion betrachtet werden. Da im gewählten *partially balanced incomplete block design* (PBIBD) (s. Seite 103) die Subkompetenzen von Konzept- und Prozesswissen nicht berücksichtigt werden konnten, wird eine empirische Testung der Zweidimensionalität vorgenommen. Die theoretischen Herleitungen lassen erwarten, dass die Kompetenz in Biologie aus den Subkompetenzen Konzept- und Prozesswissen besteht. Diese beiden Subkompetenzen sollten auf vielfältige Weisen verknüpft sein. Somit wird aufgrund der theoretischen Ausgangslage folgende Hypothese abgeleitet:

Forschungshypothese A:

Konzept- und Prozesswissen stellen zwei empirisch trennbare Kompetenzdimensionen in Biologie dar, die hoch miteinander korrelieren.

Da die Forschungshypothese auf den Zusammenhang zwischen zwei Teildimensionen innerhalb von Kompetenz in Biologie abzielt, werden Korrelationen über $r = .80$ als hoher Zusammenhang angenommen. Die theoretischen Konzeptualisierungen von *scientific literacy* weisen auf diesen hohen Zusammenhang hin (Gräber & Nentwig, 2002; Shavelson et al, 2008). Evidenz für diese Forschungshypothese wird durch einen Vergleich unterschiedlicher Modelle von Kompetenz in Biologie möglich. Das zur Fragestellung passende Modell ist ein zweidimensionales Modell mit den Faktoren Konzept- und Prozesswissen. Dieses wird gegen zwei weitere Modelle getestet. Zum einen wird es gegen ein eindimensionales Kompetenzmodell in Biologie, zum anderen gegen ein Modell mit zwei zufällig generierten Dimensionen getestet. Das Modell mit den zwei zufällig generierten Faktoren soll sicherstellen, dass im Fall eines besseren Modellfits des theoretisch begründeten zweidimensionalen Modells der Zugewinn an Modellpassung inhaltlich interpretiert werden kann und nicht nur Ursache einer (beliebig) modellierten Mehrdimensionalität ist. Die Forschungshypothese wird bestätigt, wenn das Modell Konzept- und Prozesswissen das Antwortverhalten von Schülerinnen und Schüler auf einen Kompetenztest in Biologie besser abbildet als das Zufallsmodell und das eindimensionale Modell. Bei einem Vergleich dieser drei Modelle sollte das zweidimensionale Modell Konzept- und Prozesswissen bessere Modellfit-Werte aufweisen. Zudem sollte das Zufallsmodell eine Korrelation nahe Eins zwischen den beiden Dimensionen zeigen. Wenn das Modell Konzept- und Prozesswissen eine signifikant geringere Korrelation als das Zufallsmodell aufweist, ist dies ein weiteres Indiz für die theoretisch hergeleitete Zweidimensionalität.

Erste Befunde zu den Zusammenhängen zwischen diesen beiden Wissensarten liegen in den akademischen Disziplinen Mathematik und Chemie vor (s. Seite 31 und 58). Insbesondere bei Schülerinnen und Schülern mit hohem Vorwissen konnten hohe Zusammenhänge zwischen den Subkompetenzen Konzept- und Prozesswissen gezeigt werden. In den aufgeführten Untersuchungen in der Mathematik mit Grundschülerinnen und Grundschülern liegt der Zusammenhang der Subkompetenzen zueinander bei $r = .33$ (Rittle-Johnson et al., 2001). Mit den Ergebnissen aus der Mathematik als Rückbezug auf die Naturwissenschaften sollte vorsichtig umgegangen werden, da sie zum einen ausschließlich auf Experimentalstudien mit kleinen Stichproben und wenigen Aufgaben beruhen. Zum anderen kann die Mathematik nur als ein naturwissenschaftsverwandtes Fach gesehen werden. Die in der Chemie angesiedelte Untersuchung von Klos et al. (2008) zeigte in der

Jahrgangsstufe sieben eine Korrelation von $r = .07$, in der Jahrgangsstufe 12 von $r = .68$. In der PISA-Studie 2000 korrelierte das Fachwissen mit den kognitiven Aspekten naturwissenschaftlicher Kompetenz zwischen $r = .58$ und $r = .75$. Da Prozesswissen in Biologie Sekundärprozesse des prozeduralen Wissens bzw. Problemlöseprozesse sein könnten, wird deklaratives Wissen zur Lösung der biologischen Probleme benötigt. Daher dürften beim Modellieren die beiden Dimensionen Konzept- und Prozesswissen hoch miteinander korrelieren.

Die Bearbeitung dieser Fragestellung gibt Hinweise über die interne Struktur der Kompetenz in Biologie. Mit den weiteren Fragestellungen wird diese Struktur der Kompetenz in Biologie aufgegriffen und mit verschiedenen externen Kriterien validiert.

3.2 Fragestellungen zum Zusammenhang von Kompetenz in Biologie mit verwandten kognitiven Konstrukten – Differentialpsychologische Perspektive I

Nach Forschungshypothese A wird eine zweidimensionale Struktur der Kompetenz in Biologie erwartet. Die folgenden Hypothesen werden für eine zweidimensionale Struktur der Kompetenz in Biologie expliziert, wie sie durch die Forschungshypothese A angenommen wird. Diese Struktur wird nun anhand differentialpsychologischer Fragestellungen validiert. Hierzu werden Zusammenhänge mit externen Konstrukten bearbeitet. Zusammenhänge mit verwandten Konstrukten sollten dabei höher ausfallen als mit nicht verwandten Konstrukten. In einem ersten Schritt wird untersucht, wie Kompetenz in Biologie mit den Konstrukten sprachliche Kompetenzen und kognitive Grundfähigkeit zusammenhängt. Bei beiden Konstrukten wird davon ausgegangen, dass es sich um disziplinübergreifende Fähigkeiten von Schülerinnen und Schülern handelt. Diese disziplinübergreifenden Fähigkeiten sind für die Erlangung von Kompetenz in Biologie notwendig. Zu diesem Aspekt der konvergenten und divergenten Konstruktvalidität werden daher die folgenden Hypothesen formuliert:

Forschungshypothese B1:

Sprachliche Kompetenzen und kognitive Grundfähigkeit leisten einen Beitrag zur Erklärung der beiden Dimensionen Konzept- und Prozesswissen in Biologie, erklären diese in einer Regression jedoch nicht vollständig.

Die Hypothese wird im Folgenden getrennt für die beiden Konstrukte näher erläutert.

Die theoretischen Ausführungen zu den *sprachlichen Kompetenzen* legen hohe Zusammenhänge mit Kompetenzmessungen in Biologie nahe. Für diesen hohen Zusammenhang können zwei wesentliche Gründe angeführt werden. Zum einen wird Wissen – nicht nur in den Naturwissenschaften – insbesondere über Schriftsprache erworben (Brown et al., 2005; Eschenhagen et al., 1998). Zum anderen liegen Kompetenztests im *paper-and-pencil* Format vor (s. Seite 113), sodass auch die Testadministration in hohem Maße auf Schriftsprache beruht. Bezogen auf die Lesekompetenz gehen vorangegangene Untersuchungsergebnisse stark auseinander. So zeigten die PISA Studien 2000-2009 mit Korrelationen zwischen $r = .80$ und $r = .84$ sehr starke Zusammenhänge zwischen den Lesekompetenzen und Kompetenzen in den Naturwissenschaften. Dieser besonders starke Zusammenhang kann teilweise durch die breite Erfassung beider Konstrukte erklärt werden. Die Studie von Klos et al. (2008), in der Lesekompetenz mit den verbalen KFT-Skalen operationalisiert wurde, konnte nur einen moderaten Zusammenhang von $r = .26$ mit Kompetenz in Chemie nachweisen. Allerdings erfassen diese verbalen Skalen neben Lesekompetenz auch schlussfolgerndes Denken. Für Kompetenz in Biologie liegen noch keine Zusammenhangsanalysen vor.

Die *kognitive Grundfähigkeit* ist eine notwendige Voraussetzung zur Erschließung der Welt und, um Probleme zu lösen bzw. Entscheidungen zu treffen. Sie steht daher in engem Zusammenhang zu disziplinbezogenen Kompetenzen (Rindermann, 2006; Winkelmann, 2009; Brunner, 2006) und somit zu Kompetenz in Biologie. Eine hohe kognitive Grundfähigkeit führt zu einem schnelleren und effizienteren Wissensaufbau, sodass Personen unterschiedlicher kognitiver Grundfähigkeit auch über unterschiedliche Lernfähigkeiten und unterschiedliche Lernprodukte in den akademischen Disziplinen verfügen. Für die der Biologie nahestehende Kompetenz in Chemie konnte für die Jahrgangsstufe sieben ein Zusammenhang der kognitiven Grundfähigkeit mit den naturwissenschaftlichen Arbeitsweisen von $r = .36$, für die Jahrgangsstufe 12 von $r = .46$ gezeigt werden (Klos et al., 2008). In einer weiteren Untersuchung zeigte sich für Schülerinnen und Schüler der Jahrgangsstufe 10 und 11 ein Zusammenhang von $r = .67$. Der Zusammenhang mit der Note in den Naturwissenschaften fällt geringer aus ($r = .24$) (Lau & Roeser, 2002). Die Mehrzahl der Korrelationsstudien beruht auf fachspezifischen Jahrgangsabschlussnoten. So wurden für die Naturwissenschaften Korrelationen zwischen $r = .31$ und $.36$ (Süß, 2007; Schütte et al., 2007) und für die Biologie von $r = .67$ gefunden (Lau & Roeser, 2002). Auch in der vorliegenden Arbeit wird daher von einer hohen Korrelation zwischen der kognitiven Grundfähigkeit und Kompetenz in Biologie ausgegangen.

Es ist möglich, dass die beiden übergreifenden Konstrukte unterschiedlich mit Kompetenz in Biologie zusammenhängen. Diese differentiellen Effekte werden in den beiden folgenden Hypothesen formuliert:

Forschungshypothese B2a:

Die sprachlichen Kompetenzen hängen stärker als die kognitive Grundfähigkeit mit dem Konzeptwissen der Schülerinnen und Schüler zusammen.

Forschungshypothese B2b:

Die kognitive Grundfähigkeit der Schülerinnen und Schüler hängt stärker als ihre sprachlichen Kompetenzen mit dem Prozesswissen zusammen.

Die folgende Abbildung 10 zu den Hypothesen B1 und B2 stellt die Annahmen bezogen auf die Richtung und Höhe der angenommenen Zusammenhänge der disziplinübergreifenden Konstrukte dar.

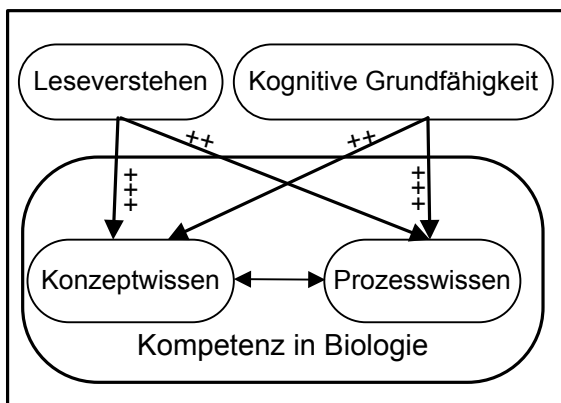


Abbildung 10: Graphische Darstellung der Forschungshypothesen B1, B2a und B2b

Anmerkungen. +++ = hohe Zusammenhänge, ++ = mittlere Zusammenhänge.

Die Anzahl der Pluszeichen an den Regressionspfaden verdeutlicht die Höhe der erwarteten Zusammenhänge. Dabei stehen die Pluszeichen nicht für ein bestimmtes Intervall von Korrelationshöhen sondern für die Höhe eines Zusammenhangs im Vergleich zur Höhe eines Zusammenhangs mit einem anderen Konstrukt. Unterschiedliche Korrelationen werden aufgrund der Gesamtschau der Zusammenhänge erwartet. So wird z. B. angenommen, dass die sprachlichen Kompetenzen höher mit dem Konzeptwissen in Biologie zusammenhängen als mit dem Prozesswissen. Das Konzeptwissen steht dem deklarativen Wissen nahe, also dem Wissen von Fakten, und das Prozesswissen dem prozeduralen Wissen. Das Prozesswissen könnte somit stärker als das Konzeptwissen mit der kognitiven Grundfähigkeit zusammenhängen, da es in hohem Maße

schlussfolgernde Aspekte des Denkens beinhaltet. Baumert et al. (2009) konnten zeigen, dass bei Kontrolle von Vorwissen der Einfluss der kognitiven Grundfähigkeit auf disziplinspezifische Leistungen von Schülerinnen und Schülern nahezu verschwindet.

Da die Kompetenz in Biologie in der vorliegenden Studie erstmalig umfassend untersucht wird, liegen für die Biologie keine theoretischen oder empirischen Grundlagen für die Höhe der relativen Zusammenhänge vor. Sie erschließen sich vielmehr indirekt aus den berichteten Ergebnissen der Mathematik und der Chemie.

3.3 Fragestellungen zum Zusammenhang von Kompetenz in Biologie mit Fachnoten und naturwissenschaftlichen Selbstkonzepten – Differentialpsychologische Perspektive II

Als weiteres Kriterium werden die Noten in unterschiedlichen Schulfächern für die konvergente und divergente Validierung herangezogen. Die Fachnote umfasst die vielfältigen Aspekte des Fachunterrichts (z. B.: mündliche und schriftliche Leistungen oder Verhalten in der Unterrichtssituation). Daher wird für Korrelationen zwischen Kompetenzen und Noten innerhalb einer Disziplin nur von einer mittleren Korrelation ausgegangen.

Forschungshypothese C1:

Die Fachnoten der Schülerinnen und Schüler stehen in einem mittleren Zusammenhang mit der erhobenen Kompetenz in Biologie.

Forschungshypothese C2:

Die Zusammenhänge der Fachnoten in den naturwissenschaftlichen Fächern Chemie und Physik mit der Kompetenz in Biologie fallen im Vergleich zu dem Zusammenhang mit den Fachnoten in Biologie geringer aus.

Forschungshypothese C3:

Die Zusammenhänge der Fachnoten in den nicht-naturwissenschaftlichen Fächern Deutsch und Mathematik mit Kompetenz in Biologie fallen wiederum geringer aus als die Zusammenhänge mit den drei naturwissenschaftlichen Fächern.

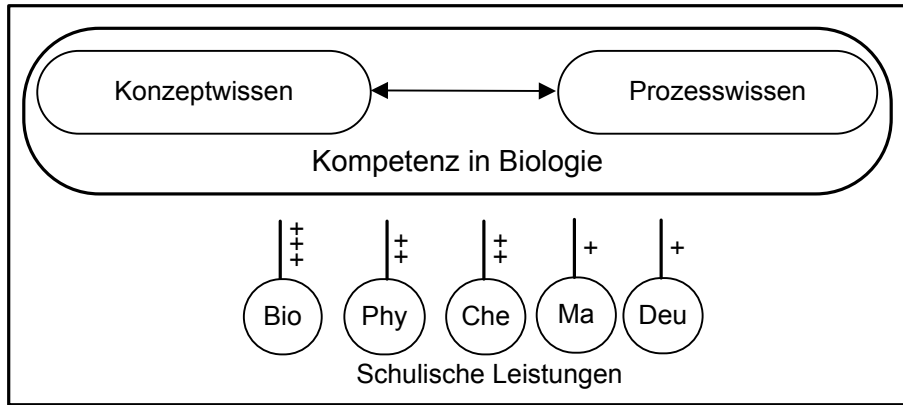


Abbildung 11: Graphische Darstellung der Forschungshypothese C

Anmerkungen. +++ = hohe Zusammenhänge, ++ = mittlere Zusammenhänge, + = geringe Zusammenhänge.

Bio = Biologie, Phy = Physik, Che = Chemie, Ma = Mathematik, Deu = Deutsch.

Ältere Untersuchungen zeigen Zusammenhänge zwischen $r = .36$ und $.66$ (Ingenkamp, 1981). Auch in neueren Untersuchungen liegen diese Zusammenhänge im Bereich mittlerer Korrelationen (z. B. Hülür et al., 2011; Lüdtke et al., 2002).

Das nomologische Netz wird zur Validierung von Kompetenz in Biologie im letzten Schritt um das Selbstkonzept in den naturwissenschaftlichen Schulfächern als externe Kriterien erweitert. Neben den disziplinübergreifenden Konstrukten und Schulnoten werden im folgenden Schritt das biologienahe und biologiefremde Selbstkonzept betrachtet. Stellvertretend für die Vielzahl von Persönlichkeitsmerkmalen wurde das weit verbreitete Merkmal Selbstkonzept in das nomologische Netz aufgenommen. Die Selbstkonzepte in den akademischen Disziplinen Biologie, Chemie und Physik werden systematisch mit Kompetenz in Biologie in Verbindung gesetzt. Die Zusammenhänge mit diesen naturwissenschaftlichen Selbstkonzepten werden in der folgenden Hypothese formuliert.

Forschungshypothese D:

Das Selbstkonzept in Biologie steht in einem höheren Zusammenhang mit der Kompetenz in Biologie als das Selbstkonzept in Chemie und Physik.

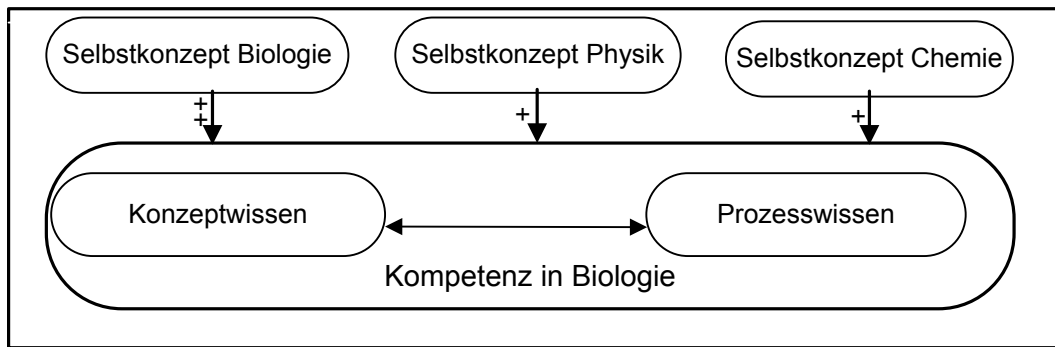


Abbildung 12: Graphische Darstellung der Forschungshypothese D

Anmerkungen. ++ = mittlere Zusammenhänge, + = geringe Zusammenhänge.

Abbildung 12 verdeutlicht, dass zwischen der Kompetenz in Biologie und dem Selbstkonzept in Biologie höhere Zusammenhänge erwartet werden als zwischen Kompetenz in Biologie und den Selbstkonzepten in Chemie sowie in Physik.

Für die Disziplinen Mathematik und Muttersprache liegen vielfältige Befunde zu den Zusammenhängen von Selbstkonzept mit akademischen Leistungen vor (z. B. Marsh & Köller, 2004; Marsh & Craven, 2006). Bezogen auf die Selbstkonzepte in den Naturwissenschaften bzw. in der Biologie kann nur auf wenige Ergebnisse zurückgegriffen werden, die ein relativ diffuses Bild zeichnen. Sie weisen sowohl auf substanzielle (Chiu, 2008; Chiu, 2012) als auch auf marginale Zusammenhänge hin (Schütte et al., 2007). Nur eine Studie bezieht alle drei naturwissenschaftlichen Selbstkonzepte ein und bestätigt das gefundene diffuse Bild (Jansen et al., in Revision).

Abschließend soll herausgestellt werden, dass für die Validierung eines Tests keine abzuarbeitenden Routinen oder gar Blaupausen vorliegen. Zudem liegen keine Grenzwerte vor, ab denen ein Testinstrument als invalide oder valide bezeichnet werden kann. Das von Cronbach und Meehl (1955) eingeführte nomologische Netzwerk liefert jedoch einen Rahmen zur Evaluation der Validität eines Tests. Die vergleichende Untersuchung von konvergenten und divergenten Validitätsaspekten mit relevanten Konstrukten gibt Aufschluss über die Verortung des Konstrukts in diesem nomologischen Netzwerk. Im Folgenden methodischen Kapitel wird darauf eingegangen, mit welcher Stichprobe und welchen Instrumenten sowie Analysen die formulierten Hypothesen innerhalb dieses Netzwerks getestet werden sollen.

Mit den skizzierten Fragestellungen und Forschungshypothesen wird für Kompetenz in Biologie erstmalig eine sehr umfangreiche Validierung vorgenommen. Die Validierung ermöglicht zum einen eine Verortung der Kompetenz in Biologie in Bezug auf die weiteren naturwissenschaftlichen Disziplinen. Zum anderen geben die Zusammenhänge Aufschluss über die vielfältigen Verknüpfungen

Kapitel 3

zu weiteren Wissensbeständen der Schülerinnen und Schüler. Diese Verknüpfungen sollten in der schulischen und außerschulischen Praxis berücksichtigt werden.

4 Methode

Die formulierten Fragestellungen werden anhand eines umfangreichen Datensatzes aus dem Projekt ESNaS zur Erfassung der in den Bildungsstandards beschriebenen Kompetenzen in den Naturwissenschaften für die Kompetenzbereiche Erkenntnisgewinnung und Fachwissen untersucht. Die eingesetzten Aufgaben sind in besonderem Maße für die Beantwortung der Fragestellungen geeignet, da im Rahmen dieses Projekts erstmalig eine hohe Anzahl an Testaufgaben zu den beiden Teilbereichen Konzept- und Prozesswissen in den Naturwissenschaften entwickelt wurde. Eine einzelne Testaufgabe kann dabei entweder dem Konzept- oder dem Prozesswissen zugeordnet werden. Damit ermöglichen die Aufgaben differenzierte Analysen zu disziplinbezogenem Prozesswissen und disziplinbezogenem Konzeptwissen. Die Prämissen im Aufgabenentwicklungsprozess, die Fülle der Testaufgaben sowie das Testdesign lassen zuverlässige Aussagen auch für eine einzelne naturwissenschaftliche Disziplin zu, für die vorliegende Arbeit die Biologie. Zunächst werden die Zielpopulation und die daraus resultierende Stichprobe beschrieben. Nach der Operationalisierung der verwendeten Konstrukte wird auf die Durchführung der Studie und das Multi-Matrix Design sowie die statistischen Analysen eingegangen.

4.1 Grundgesamtheit und Stichprobenziehung

Die Zielpopulation sind Schülerinnen und Schüler in Deutschland, die mindestens den Mittleren Schulabschluss (MSA) anstreben. Mit dem MSA beendeten im Jahr 2010 41 % der Schülerinnen und Schüler ihre Schullaufbahn (KMK, 2011). Der MSA wird am Ende der Jahrgangsstufe 10 absolviert. Die Bildungsstandards für die Sekundarstufe I sind für den MSA konzipiert und zielen auf das Ende der Jahrgangsstufe 10. Bis zu diesem Zeitpunkt kann das gesamte von den Schülerinnen und Schülern erreichte Leistungsspektrum untersucht werden, bevor ein Großteil der Population auf den Arbeitsmarkt entlassen wird.

Die dieser Untersuchung zu Grunde liegenden Daten sind Teil einer groß angelegten Studie zur Erfassung der in den Bildungsstandards beschriebenen Kompetenzen in den naturwissenschaftlichen Fächern Biologie, Physik und Chemie. Für die Analysen werden nur die Schülerinnen und Schüler berücksichtigt, denen Biologieaufgaben vorgelegt wurden. Daher wird zunächst die Generierung der Gesamtstichprobe Naturwissenschaften beschrieben, worauf die Beschreibung der Schritte zur Festlegung der Substichprobe Biologie folgt.

Die Ziehung der Gesamtstichprobe wurde in drei Schritten vollzogen. In einem ersten Schritt wurden die acht Bundesländer Nordrhein-Westfalen, Schleswig-Holstein, Berlin, Sachsen, Bayern, Baden-Württemberg, Hessen und Thüringen ausgewählt. Bei der Auswahl wurde die Heterogenität des föderalen Bildungssystems berücksichtigt, indem z. B. alte und neue Bundesländer sowie Stadtstaaten Eingang fanden. In einem nächsten Schritt wurden je Bundesland per Zufall 20 Schulen gezogen, an denen der MSA erworben werden kann. Der MSA wird vorrangig an Gymnasien, Gesamt- und Realschulen sowie Schulen mit mehreren Bildungsgängen (MBG) erworben; Hauptschulen wurden daher nicht in die Schulstichprobe aufgenommen. Somit besteht die angestrebte Schulstichprobe aus 160 Schulen der Sekundarstufe I. Nicht allen zufällig gezogenen Schulen war eine Teilnahme im Testzeitraum möglich, daher mussten für 6 % der Schulen Ersatzschulen gezogen werden. Trotz dieser Nachziehungen konnte aus organisatorischen bzw. administrativen Gründen für zwei Schulen kein Ersatz gefunden werden, sodass die Testung in 158 Schulen realisiert werden konnte. Die Teilnahmequote beträgt also auf Schulebene 99 %. In einem letzten Schritt wurden in jeder dieser 20 Schulen zufällig zwei vollständige 10. Klassen gezogen. Somit besteht die angestrebte Stichprobe auf Klassenebene aus 320 Einheiten, von denen aufgrund der Ausfallquote auf der höher gelegenen Schulebene nur 316 möglich waren. Da 17 Schulen nur einen einzügigen 10. Jahrgang hatten, konnten nur 299 Klassen getestet werden. Dies entspricht 94 % der angestrebten Klassen.

In jeder dieser Klassen wurden Biologie-, Chemie- und Physiktesthefte sowie Mischformen eingesetzt (s. Seite 104). Von den 6 084 Schülerinnen und Schülern, die an der Untersuchung teilgenommen haben, erhielten 3 156 Schülerinnen und Schüler entweder Biologie- oder Mischtesthefte, die Biologieaufgaben beinhalteten. Das Verfahren führte zu einer durchschnittlichen Schülerteilnahmequote von 11 Schülerinnen und Schülern pro Klasse; der Median pro Klasse liegt ebenfalls bei 11 Schülerinnen und Schülern.

Die Beschreibung der Stichprobenziehung zeigt, dass es sich um eine per Zufall generierte, nicht-stratifizierte und natürliche Klumpenstichprobe zu einem Messzeitpunkt handelt. Bei einer Klumpenstichprobe ist die Grundlage der Ziehung nicht das einzelne Individuum sondern eine soziale Einheit. Die gezogenen Schulen und Klassen sind solche bestehenden sozialen Einheiten. Häufig ist die Merkmalsvariabilität innerhalb dieser Einheiten geringer als die von Personen zwischen den Einheiten. Als Maß für die Homogenität dieser Einheiten kann die Intraklassenkorrelation berechnet werden (Lüdtke, Trautwein, Kunter & Baumert, 2006). Diese beträgt für den eindimensional skalierten Personenparameter *weighted likelihood estimator* (wle) aller Biologieaufgaben auf Schulebene $ICC(2) = .84$, auf Klassenebene $ICC(2) = .75$. Diese Werte sprechen dafür, dass ein substantieller Anteil der Gesamtvarianz durch die Zugehörigkeit zu Klassen und Schulen erklärt

werden kann. Im Abschnitt statistische Analysen wird darauf eingegangen, wie dieser Befund bei den Berechnungen berücksichtigt wurde.

Abschließend muss festgehalten werden, dass die vorliegende Stichprobe nicht repräsentativ für Deutschland ist. Die Ausführungen zur Ziehung der Stichprobe zeigen jedoch, dass die für den MSA relevanten Schulformen enthalten sind. Somit ist die Generalisierbarkeit der Ergebnisse von Zusammenhangsanalysen zulässig. Es kann davon ausgegangen werden, dass Zusammenhänge, die in der vorliegenden Stichprobe auftreten, auch in der Zielpopulation wiedergefunden werden.

Tabelle 15: Vergleich der Biologie-, der Gesamtstichprobe und der eliminierten Fälle (Nicht-Biologie) anhand des Alters und des Geschlechts (Geschlecht in Prozent)

Kriterium	Biologie	Nicht-Biologie	Gesamtstichprobe
Geschlecht	w = 51 / m = 49	w = 53 / m = 47	w = 51 / m = 49
Alter	15.54	15.50	15.50

Anmerkungen. w = weiblich, m = männlich.

Des Weiteren wird in der Tabelle 15 sichtbar, dass die Selektion der Schülerinnen und Schüler, die Biologie-Testhefte erhalten haben, bezogen auf Geschlecht und Alter zu keiner Verzerrung der Stichprobe geführt hat. Bezüglich des Geschlechts und des Alters unterscheidet sich die Gesamtstichprobe Naturwissenschaften nicht von der Substichprobe Biologie. Die Schülerinnen und Schüler verteilen sich wie folgt auf die berücksichtigten Schulformen: 47 % Gymnasium, 35 % Realschule, 10 % Schulen mit mehreren Bildungsgängen, 7 % Gesamtschulen (Abweichungen von 100 % ergeben sich durch Rundungsungenauigkeiten).

4.2 Durchführung der Studie

Da die Jahrgangsstufe 10 besonders im zweiten Schulhalbjahr durch eine Vielzahl an Prüfungsbelastungen gekennzeichnet ist, wurde die Untersuchung zu Beginn des Schuljahres durchgeführt. Die Testungen fanden zwischen dem 28. September und dem 16. Oktober 2009 statt. Mit der Organisation und Durchführung der Testung in den acht Bundesländern wurde das IEA Data Processing and Research Center (DPC) betraut. Jeweils zwei geschulte Testleiterinnen bzw. Testleiter führten die Testung an einer ihnen zugewiesenen Schule durch. Sie bekamen das Testmaterial für ihre Schule zugesandt und führten in Eigenverantwortung die Testung durch. Um gleiche Testbedingungen an allen teilnehmenden Schulen gewährleisten zu können, mussten die Testleiterinnen und Testleiter genau die Vorgaben eines eigens für die Studie erstellten

Testleiterskripts befolgen und wurden vor der Testung diesbezüglich geschult. Durch das Vorlesen dieses Testleiterskripts wurden der zeitliche und administrative Ablauf und die Anweisungen der Testleiterinnen und Testleiter an die Klasse festgelegt.

Die Durchführung der Testung nahm ca. dreieinhalb Stunden in Anspruch. Die Aufteilung der Zeit auf die unterschiedlichen Testteile kann Tabelle 16 entnommen werden.

Tabelle 16: Zeitlicher Ablauf der Testung

Minuten	Testabschnitt
60	Bearbeitung von Testaufgaben – Teil 1
15	Pause
60	Bearbeitung von Testaufgaben – Teil 2
15	Pause
20	Bearbeitung des KFT und des Lückentests
40	Bearbeitung des Fragebogens

Im ersten Teil der Testung bearbeiteten die Schülerinnen und Schüler in zwei 60-minütigen Abschnitten die naturwissenschaftlichen Testaufgaben. Da in diesem Zeitraum nicht alle Aufgaben von allen Schülerinnen und Schülern bearbeitet werden konnten, wurde ein Multi-Matrix Design angewendet (Gonzalez & Rutkowski, 2010). Bei diesem Design bekommt eine einzelne Schülerin bzw. ein einzelner Schüler nur einen Teil der Aufgaben vorgelegt. Die Aufgaben wurden in 20-minütige Zeitabschnitte, so genannte Blöcke, eingeteilt. Die Testzeit von 120 Minuten konnte somit in jedem Testheft mit sechs dieser 45 20-minütigen Blöcke aufgefüllt werden. Da nicht alle Blöcke in jedem Testheft eingesetzt werden konnten, musste ein *incomplete block design* (IBD) verwendet werden. Um in einem solchen Design Positionseffekte zu kontrollieren, wird versucht, jeden Block einmal an jeder Blockposition zu platzieren. Ein solches Design wird *balanced incomplete block design* (BIBD) genannt (Frey, Hartig & Rupp, 2009). Dabei müssen die Bedingungen erfüllt werden, dass (a) Jeder Block mindestens einmal in einem Testheft erscheint, (b) jeder Block über alle Testhefte in gleicher Anzahl auftritt, (c) jedes Testheft die gleiche Länge und die gleiche Anzahl von Blöcken enthält und (d) jedes Paar von aufeinanderfolgenden Blöcken in gleicher Anzahl über die Testhefte auftritt (Frey et al., 2009). Die Umsetzung eines solchen Designs hätte zu einer unverträglich hohen Anzahl an Testheften geführt. Daher konnte jeder Block nur an drei von sechs möglichen Positionen gesetzt werden. Im Folgenden soll dieses Design als PBIBD bezeichnet werden. Jeder Block lief also entweder zwei Mal in der zweiten Testhälfte und einmal in der ersten Testhälfte oder *vice versa*. Diese Abänderung des BIBD führte dazu, dass über alle Testhefte nicht alle Blockpaare in der gleichen

Anzahl vorkommen. Die vierte Bedingung wurde zu Gunsten der Verhinderung von Blockpositionseffekten nicht eingehalten. Das Testdesign kann daher nur als *partially balanced* bezeichnet werden. Es muss jedoch angemerkt werden, dass auch das gewählte Design für potenzielle Positionseffekte auf Seiten der Testaufgaben ebenfalls nicht vollständig, sondern nur teilweise kontrolliert ist. Über dieses PBIBD können alle Testaufgaben miteinander verknüpft werden. Beispiele für die unterschiedliche Positionierung des gleichen Blocks in den verschiedenen Testheften sind in Tabelle 17 dargestellt. So läuft der 20-Minuten Block BioFW14 (Biologie Fachwissen, Block 14) in Testheft vier an der ersten Position, in Testheft fünf an der vierten Position. In Testheft 71, welches nicht in der Tabelle abgebildet ist, steht dieser Block an der sechsten Position. Der Fokus des Testdesigns lag auf der Verlinkung der beiden Teilkompetenzen Prozess- und Konzeptwissen.

Tabelle 17: Auszug aus dem Testdesign der Studie – Biologietesthefte

Position	1	2	3	4	5	6
Testheft04	BioFw14 ¹	BioFw11	BioFw07	BioEg10 ⁷	BioEg17 ⁴	BioFw18
Testheft05	BioFw15	BioFw01 ²	BioEg11	BioFw14 ¹	BioEg08	BioFw10 ⁸
Testheft06	BioFw12	BioFw20 ³	BioEg01	BioEg12	BioFw01 ²	BioFw19 ⁵
Testheft07	BioEg02	BioFw17 ⁴	BioEg06	BioFw22	BioFw20 ³	BioFw21 ⁹
Testheft08	BioFw03	BioEg15	BioFw17 ⁶	BioFw10 ⁸	BioFw09	BioEg10 ⁷
Testheft09	BioFw04	BioFw19 ⁵	BioEg05	BioFw21 ⁹	BioFw02	BioEg17 ⁶

Anmerkungen. Bio, Che, Phy = akademische Disziplin; Fw, Eg = Kompetenzbereich. Gleiche Blöcke an unterschiedlichen Stellen sind mit den gleichen Ziffern indiziert.

Da die Studie für die Naturwissenschaften angelegt ist, hat nur ein Teil der Schülerinnen und Schüler einer Klasse Biologie-Testhefte vorgelegt bekommen. Insgesamt wurden 73 verschiedene Versionen der Testhefte mit naturwissenschaftlichen Aufgaben von der gesamten Stichprobe beantwortet. Von diesen waren 22 Versionen komplette Biologie-Testhefte, 16 Versionen bestanden aus einem Biologie und einem Chemie-, einem Physik- oder einem Chemie/Physikteil. Im 80-minütigen zweiten Teil wurde von allen Schülerinnen und Schülern der KFT-Test und zwei Lückentests (s. Seite 116) bearbeitet, sowie ein Fragebogen zu ihren Hintergrundmerkmalen und Selbsteinschätzungen bezüglich der Naturwissenschaften ausgefüllt.

Das Testmaterial wurde unmittelbar nach der Testung von den Testleiterinnen und Testleitern an das DPC zurückgesandt. Die Antworten im *multiple-choice* (MC-) Format (s. Seite: 107) wurden maschinell eingelesen, die Aufgaben mit offenem Antwortformat wurden von geschulten Kodiererinnen und Kodierern anhand von Kodieranweisungen bewertet. Die hierfür notwendigen Kodieranweisungen wurden von durch die Bundesländer benannten Lehrkräften erstellt. Nach der Durchführung der Testung wurden an vier Terminen Kodiererschulungen mit den am Aufgabenentwicklungsprozess beteiligten Fachdidaktikerinnen und Fachdidaktikern durchgeführt, um eine hohe Interraterreliabilität zu erreichen. In diesen Schulungen besprachen die Kodiererinnen und Kodierer mit den Biologiedidakterinnen und -didaktikern jene Antworten, die unterschiedlich bewertet wurden. Mit diesen Ergebnissen wurden die Kodieranweisungen im Nachhinein noch einmal überarbeitet. Die Schülerantworten wurden daraufhin in der finalen Kodierung jeweils einer Kodiererin bzw. einem Kodierer von 16 zur Bewertung vorgelegt.

4.3 Operationalisierung der Konstrukte

Die Beschreibung der verwendeten Instrumente erfolgt in zwei Teilabschnitten. Zunächst werden die neu entwickelten Testaufgaben zu Konzept- und Prozesswissen beschrieben. Daraufhin folgt eine knappe Beschreibung der Instrumente zur Validierung dieser Testaufgaben. Diese wurden bereits in deutschlandweiten und internationalen Schulleistungsuntersuchungen eingesetzt.

4.3.1 Testaufgaben zum Konzept- und Prozesswissen in Biologie

Im Rahmen des Projekts ESNaS wurden 90 einleitende biologische Texte und dazugehörige Testaufgaben entwickelt. Pro Einführungstext müssen die Schülerinnen und Schüler mehrere Testaufgaben bearbeiten. Jede dieser Aufgaben kann entweder dem Prozess- oder dem Konzeptwissen zugeordnet werden. Für die Schülerinnen und Schüler sind alle relevanten Informationen zur Lösung der Aufgaben in dem einleitenden biologischen Text vorgegeben. Dieses Vorgehen wurde gewählt, da aufgrund der inhaltlich heterogenen Curricula zwischen den Bundesländern kaum naturwissenschaftliche Sachverhalte auszumachen sind, die alle Schülerinnen und Schüler bis zur Jahrgangsstufe 10 gleichermaßen gelernt haben. Somit würden Ergebnisse auf Grundlage der Aufgaben zum Faktenwissen in erster Linie das Behandeln unterschiedlicher Themengebiete in den einzelnen Bundesländern abbilden. Um abzusichern, dass die Aufgaben naturwissenschaftliche Kompetenz testen und nicht vorrangig sprachliche Kompetenzen anhand naturwissenschaftlicher Inhalte, wurden zusätzlich sprachliche Kompetenzen erhoben (s. Seite 116).

Einer allgemeinen Beschreibung der Entwicklung der Testaufgaben folgt eine Vorstellung der Aufgaben zu den beiden Wissensarten.

4.3.1.1 Entwicklung der Testaufgaben zum Konzept- und Prozesswissen

Im Zeitraum zwischen Winter 2008 und Frühjahr 2009 wurden von neun aus den Bundesländern benannten Lehrkräften Aufgaben zu den Kompetenzbereichen Fachwissen und Erkenntnisgewinnung der Bildungsstandards im Fach Biologie entwickelt.

Als gemeinsame Grundlage für die Aufgabenentwicklung in den naturwissenschaftlichen Fächern Chemie, Physik und Biologie wurden von Fachdidaktikerinnen und -didaktikern umfangreiche, fachübergreifende Konstruktionsanleitungen (für Fachwissen, noch nicht veröffentlicht; für Erkenntnisgewinnung, noch nicht veröffentlicht⁶) und ein Modell der schwierigkeiterzeugenden Merkmale (Kauertz, 2008) erarbeitet. Zudem übernahmen die Fachdidaktikerinnen und -didaktiker die Koordination der Aufgabenentwicklung in ihrer jeweiligen akademischen Disziplin (Biologie: Prof. Jürgen Mayer, Dipl.-Biol. Nicole Wellnitz und Dipl.-Psych. Stefan Hartmann). Da die Lehrkräfte in ihren Bundesländern angesiedelt sind, wurden zur Sicherung der Qualität der länderübergreifenden Aufgaben folgende Strukturen geschaffen:

- halbjährliche fachübergreifende dreitägige Tagungen, die den intra- und interdisziplinären Austausch ermöglichten,
- regelmäßige, ca. vierteljährliche eintägige Treffen in den Fachgruppen sowie
- Austauschstrukturen zwischen zwei Lehrkräften innerhalb einer Fachgruppe,
- Betreuung durch eine fachdidaktische Beraterin des Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Frau Prof. Harms und
- ein umfangreicher fachdidaktischer, psychometrischer und linguistischer Bewertungsprozess.

Die fachdidaktische Prüfung fokussierte die inhaltliche Relevanz und Richtigkeit der Testaufgaben, die psychometrische Prüfung die Eignung als Testaufgabe und die linguistische Überprüfung die sprachlichen Aspekte. Zur weiteren Qualitätssicherung wurden die Aufgaben von den Biologiedidaktikerinnen und -didaktikern an kleinen Schülerstichproben erprobt und auf Grund der Ergebnisse einer weiteren Überarbeitung unterzogen. Der beschriebene Aufgabenentwicklungsprozess ist in der folgenden Abbildung 13 verdeutlicht.

⁶ Die Veröffentlichung zum Kompetenzbereich Fachwissen ist momentan in der Bearbeitung.

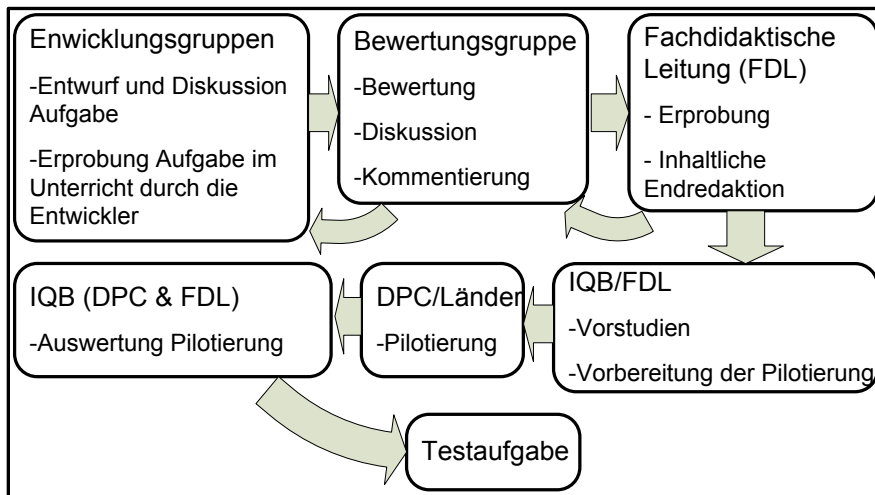


Abbildung 13: Darstellung des Aufgabenentwicklungsprozesses

Anmerkungen. IQB = Institut zur Qualitätsentwicklung im Bildungswesen, DPC = Data Processing and Research Center.

Die Merkmale und Strukturen der finalen, eingesetzten Testaufgaben werden im Folgenden näher beschrieben.

4.3.1.2 Beschreibung der Aufgabenmerkmale

Die 90 entwickelten Aufgabeneinheiten weisen die gleiche Struktur auf. Nach einem einleitenden biologischen Text, der in die Thematik und die Problemstellung einführt, folgen durchschnittlich drei meist frei kombinierbare Testaufgaben. Diese 270 Testaufgaben liegen in den unterschiedlichen Antwortformaten MC (51 %), Kurzantwort (13 %) oder erweiterte Antwort (36 %) vor. Im MC-Format muss eine von vier alternativen Antworten ausgewählt werden. Bei der Kurzantwort müssen die Schülerinnen und Schüler meist mit einem Wort oder einer Zahl auf die Fragestellung antworten. Bei dem erweiterten Antwortformat müssen sie eine längere Antwort von mindestens einem Satz schreiben. Die Fülle der Testaufgaben ermöglicht detaillierte und abgesicherte Analysen für das Konzept- und Prozesswissen in den einzelnen Disziplinen Biologie, Physik und Chemie. Um Testaufgaben zu generieren, die das gesamte Schwierigkeitsspektrum abdecken, wurde ein Aufgabenentwicklungsmodell entwickelt (Kauertz, 2008). Dieses Modell unterscheidet zwei schwierigkeitszeugende Merkmale von Testaufgaben; zum einen die kognitiven Prozesse (a) Reproduzieren, (b) Selektieren, (c) Organisieren und (d) Integrieren im Sinne der Anforderungen an die Schülerin bzw. den Schüler, zum anderen die Komplexität einer Aufgabe mit den Abstufungen (a) ein Fakt, (b) zwei Fakten, (c) ein Zusammenhang, (d) zwei Zusammenhänge und (e) übergeordnetes Konzept. Laut diesem Modell müssen die Schülerinnen und Schüler bei den leichtesten Testaufgaben aus einem gegebenen naturwissenschaftlichen Text einen Fakt reproduzieren; bei der schwersten Testaufgabe muss ein übergeordnetes Konzept integriert werden.

Das Aufgabenentwicklungsmodell mit seinen a priori angenommenen Schwierigkeiten erlaubt die Entwicklung von Testaufgaben unterschiedlicher Schwierigkeit. Die Schwierigkeitsstreuungen werden auf einer Logit-Skala berichtet (s. Seite 120). Im Folgenden werden die Testaufgaben zu den beiden Subkompetenzen Konzept- und Prozesswissen detailliert vorgestellt.

Beschreibung der Testaufgaben zum Konzeptwissen

Das Konzeptwissen in Biologie wird über die Testaufgaben zu dem Kompetenzbereich Fachwissen der Bildungsstandards für das Fach Biologie abgebildet. Konzeptwissen in Biologie ist – angelehnt an die Definition von Kompetenz in Biologie (s. Seite 43) – in die Basiskonzepte System (32 % der Testaufgaben), Struktur und Funktion (44 %) sowie Entwicklung (25 %) gegliedert (Abweichungen von 100 % ergeben sich durch Rundungsungenauigkeiten). Die Aufgaben zum Kompetenzbereich Fachwissen bilden das biologische Inhaltswissen der Schülerinnen und Schüler somit breitgefächert ab. Die benannten Lehrkräfte entwickelten 52 einleitende biologische Fachwissenstexte und 117 zugehörige Testaufgaben. Von diesen lagen 51 % im MC-Format, 20 % als Kurzantwort und 29 % als erweiterte Antwort vor. Wie bereits erwähnt, sind die Informationen zur Lösung der Testaufgabe im einleitenden Text gegeben. Durch diese Konstruktion der Aufgaben wurde verhindert, dass die Schülerinnen und Schüler zur Bearbeitung einer Aufgabe des Konzeptwissens Prozesswissen benötigen und *vice versa*. In der folgenden Abbildung 14 wird eine Beispielaufgabe aus dem Kompetenzbereich Fachwissen wiedergegeben.

Vergiftete Pfeile

Naturwissenschaftlicher Einführungstext:

Seit Hunderten von Jahren nutzen die südamerikanischen Ureinwohner das Pfeilgift „Curare“ bei der Jagd. Dieses Gift aus dem Extrakt von Pflanzen kann Tiere innerhalb kürzester Zeit völlig bewegungsunfähig machen und sogar töten.

Möglich ist dies, da „Curare“ die Übertragung von Erregungen zwischen Nervenzellen und Muskelzellen verhindert. Damit lähmt es die Muskeln, die das Tier zum Bewegen des Körpers braucht. Auch die Atemmuskulatur ist betroffen. Das Gift hat dagegen keine Wirkung auf den Herzmuskel. Die Wirkung von „Curare“ im Organismus lässt nach einiger Zeit wieder nach.

Beispielaufgabe:

Im 18. Jahrhundert verabreichte ein Wissenschaftler Säugetieren eine tödliche Dosis „Curare“. Er schaffte es trotzdem, sie am Leben zu erhalten.

Beschreibe, welche lebensrettende Maßnahme er durchführen musste, um den Tod der Tiere zu verhindern.

--

Abbildung 14: Beispielaufgabe zum Konzeptwissen in Biologie

Diese Aufgabe im erweiterten Antwortformat behandelt das System eines tierischen Körpers. Um sie zu lösen, müssen nach dem Aufgabenentwicklungsmodell die zwei Zusammenhänge Wirkung des Nervengifts Curare und Wirkung unterschiedlicher lebensrettender Maßnahmen auf die tierischen Muskelsysteme integriert werden. Erst durch die Integration dieser beiden Zusammenhänge kann die Schülerin bzw. der Schüler zu der Lösung kommen, dass der Wissenschaftler die Säugetiere über künstliche Beatmung am Leben halten konnte.

Wie bereits erwähnt, konnten den Schülerinnen und Schülern nicht alle Testaufgaben vorgelegt werden. Die Aufgaben zum Konzeptwissen wurden im Durchschnitt von 254 Schülerinnen und Schüler beantwortet ($\text{Max}_{\text{Anzahl Personen}} = 503$; $\text{Min}_{\text{Anzahl Personen}} = 242$, $\tilde{x}_{\text{Anzahl Personen}} = 252$). Die Schülerinnen und Schüler erhielten für jede gelöste Aufgabe einen Punkt, für jede falsche Aufgabe null Punkte. Die Daten für die Auswertung liegen somit dichotom vor. Der folgenden Abbildung 15 ist zu entnehmen, wie viele Schülerinnen und Schüler welche Anzahl an Testaufgaben bearbeitet haben.

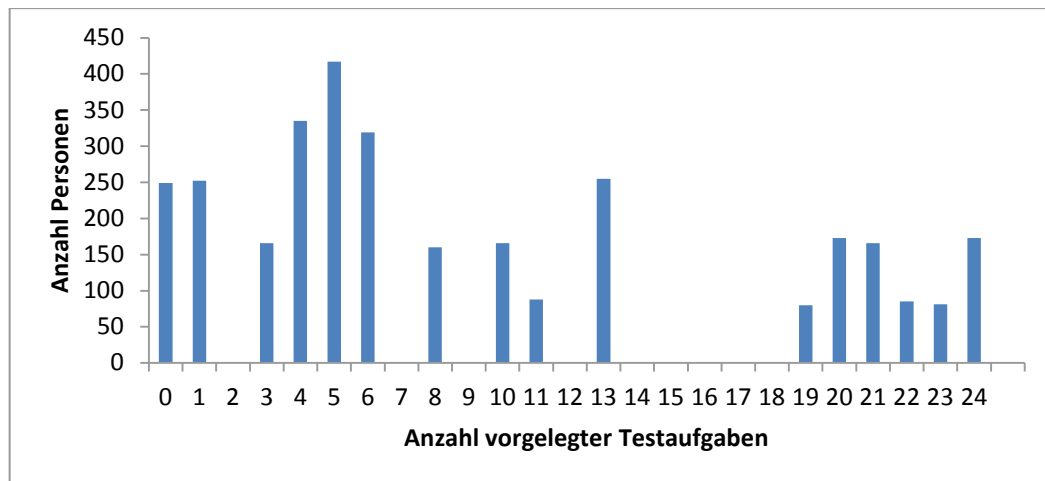


Abbildung 15: Anzahl vorgelegter Testaufgaben für den Kompetenzbereich Fachwissen

Von den 3 165 Schülerinnen und Schülern haben 2 916 Fachwissensaufgaben bearbeitet. Von diesen haben 34 % der Schülerinnen und Schüler vier bis sechs Testaufgaben vorgelegt bekommen, die Spanne reicht bis zu 24 vorgelegten Testaufgaben. In Abbildung 16 ist die Schwierigkeitsverteilung der Testaufgaben zum Kompetenzbereich Fachwissen auf der durch ein *one paramater logistic* (1PL) Modell definierten Logit-Skala abgetragen. Das zu Grunde liegende Messmodell wird im Abschnitt statistische Analysen eingehend beschrieben (s. Seite 120).

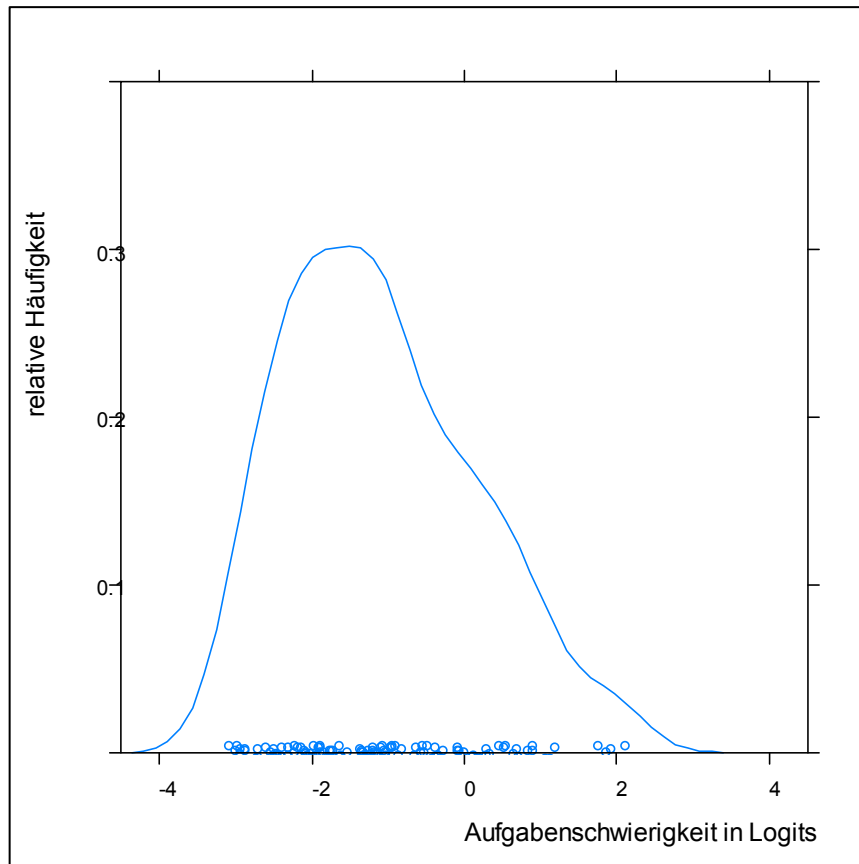


Abbildung 16: Schwierigkeitsverteilung der Testaufgaben zu Konzeptwissen in der Jahrgangsstufe 10
Anmerkungen. Jeder Punkt steht für eine Testaufgabe.

Die mittlere Schwierigkeit der Testaufgaben beträgt $M_b = -1.10$ ($SD_b = 1.23$, $Max_b = 2.12$, $Min_b = -3.08$). Die modellbasierte Reliabilität der Skala Biologie Konzeptwissen liegt bei McDonald's $\omega = .98$ (McDonald, 1999)

Beschreibung der Testaufgaben zum Prozesswissen

Die Herangehensweisen zur Testung von Prozesswissen in den Naturwissenschaften können grob in praktische und schriftliche Testformen eingeteilt werden. Für das Prozesswissen in den Naturwissenschaften werden in Schulleistungsstudien weitestgehend schriftliche Tests verwendet, es sind jedoch verschiedene Herangehensweisen möglich. Die Vor- und Nachteile dieser Herangehensweisen werden im Folgenden diskutiert. Im Anschluss wird die Operationalisierung analog zum Konzeptwissen beschrieben.

Beide Testformen unterscheiden sich voneinander (Doran & Tamir, 1992). Während praktische Tests – so genannte *hands-on* Tests – den Vorteil haben, dass sie den realen Prozess des Wissenserwerbs nachstellen, können in schriftlichen Tests zusätzlich zu diesen Arbeitsweisen auch

Denkweisen in die Testung mit einbezogen werden. Mit einem praktischen Test können Manipulationen von Apparaten und Beobachtungen, mit einem schriftlichen Test Verwendung von Graphen, Untersuchungsplanung und Dateninterpretation abgebildet werden.

Diese Unterschiede bringen mit sich, dass durch die Verwendung unterschiedlicher Herangehensweisen in der Testung von Prozesswissen in der Biologie nur bestimmte Aspekte dieses Prozesswissens gemessen werden können. In der Durchführung erweisen sich praktische Tests als aufwändig, zudem ist ihre Reliabilität häufig nicht zufriedenstellend (Baumert et al., 1997; Solano-Flores & Shavelson, 1997). Andererseits zeigte sich, dass insbesondere bei unerfahrenen Schülerinnen und Schülern eine Diskrepanz zwischen geschriebenen Berichten und der tatsächlichen Durchführung einer naturwissenschaftlichen Untersuchung vorzufinden ist (Harlow & Jones, 2004; Roberts & Gott, 2004). So konnte gezeigt werden, dass Schülerinnen und Schüler, die auf dem Papier sinnvolle Planungen von naturwissenschaftlichen Untersuchungen vornehmen, diese nicht zwangsläufig in die Tat umsetzen können (Gott & Duggan, 1999; Roberts & Gott, 2004; Tamir et al. 1992). In der *First International Science Study* (FISS) wurden drei verschiedene Tests zur Überprüfung naturwissenschaftlicher Kompetenz eingesetzt: ein schriftlicher theoretischer Test, ein schriftlicher Test zum praktischen Aspekt der Naturwissenschaften und ein praktischer Test. Der praktische Test korrelierte in der Population der 14-Jährigen mit dem theoretischen Test zu $r = .39$, mit dem schriftlichen praktischen Teil zu $r = .45$. Die beiden schriftlichen Tests korrelierten hingegen zu $r = .73$. Dieser Befund wurde in der *Second International Science Study* (SISS) repliziert (Tamir et al., 1992). In NAEP wurden bei der Erhebung der Fähigkeiten in den Naturwissenschaften im Jahre 2009 die Ergebnisse der praktischen Tests getrennt von den schriftlichen Tests berichtet (NCES, 2011). Schülerleistungen in praktischen Tests korrelieren also nur schwach mit Schülerleistungen in nicht-praktischen Tests (Hammann et al., 2007).

Den Reliabilitätsproblemen des praktischen Testens und den Problemen des Aufwands wurde mit der Untersuchung alternativer Testmethoden, wie z. B. Computersimulationen oder die Möglichkeit des freien Antwortens, begegnet (Shavelson, Baxter & Pine, 1991; Solano-Flores & Shavelson, 1997). Zwischen diesen Alternativmethoden und den praktischen Tests konnten nur moderate Korrelationen von $r = .36$ bis $r = .41$ gezeigt werden (Shavelson et al., 2008; Tamir et al., 1992). Einschränkend muss darauf hingewiesen werden, dass praktische Tests oft in Gruppen, Alternativtests als Einzeltests durchgeführt werden. Die moderaten Korrelationen zeigen, dass *paper-and-pencil* Tests und computerbasierte Erhebungen von Prozesswissen unterschiedliche Aspekte des Prozesswissens messen und keine Ersatztests für praktische Tests darstellen.

Die Durchführung von sogenannten Ersatztests sowie die Durchführung von Experimenten an Computern könnten in einer Stichprobe des vorliegenden Umfangs nur mit hohem Mitteleinsatz

durchgeführt werden. So wird auch in den gängigen (deutschen) Schulleistungsuntersuchungen (s. PISA oder TIMSS) eine schriftliche Testung gewählt. Weltweit wird nur in zwei *large-scale assessment* Studien zur Erhebung der naturwissenschaftlichen Kompetenz eine schriftliche mit einer praktischen Testung kombiniert, in der FIS- und SIS – Studie sowie im NAEP in den USA (Tamir & Doran, 1992; Winick et al., 2008;). Praktische Testungen sollten daher eher in schulischen Zusammenhängen genutzt werden (Tamir et al., 1992). In der PISA-Studie 2012 werden die Testaufgaben zum Problemlösen computerbasiert getestet (Martin, 2011).

Für die vorliegende Untersuchung wurde ein schriftlicher Test zum Prozesswissen gewählt. Mit dieser Entscheidung geht allerdings einher, dass nicht sämtliche Aspekte des Prozesswissens in Biologie erhoben werden können. Bezogen auf die manuellen Fertigkeiten ist offensichtlich, dass Teile von Kompetenzen (und von Bildungsstandards) nicht oder nur schwer messbar sind (Köller, 2008). Das Prozesswissen wird in der vorliegenden Arbeit mit 38 einleitenden biologischen Texten und 153 dazugehörigen Testaufgaben zum Kompetenzbereich Erkenntnisgewinnung der Bildungsstandards für das Fach Biologie abgebildet. Die drei großen Themengebiete der Erkenntnisgewinnung sind laut Definition der Kompetenz in Biologie (s. Seite 43) die wissenschaftstheoretische Reflexion (14 % der Testaufgaben), die naturwissenschaftliche Modellbildung (16 %) und die naturwissenschaftliche Untersuchung (69 %). Die wissenschaftstheoretische Reflexion umfasst das Verständnis der Eigenschaften naturwissenschaftlichen Wissens sowie der Bedingungen ihrer Entwicklung. Sie zielt damit auf eine Metaebene, also auf die Ebene der Reflexion über den naturwissenschaftlichen Erkenntnisgewinnungsprozess und die Charakteristika der Biologie als wissenschaftliche Disziplin ab. Die naturwissenschaftliche Modellbildung umfasst das Wissen um die Funktionalität und die adäquate Anwendung von Modellen sowie das Wissen um deren Grenzen. Der Bereich naturwissenschaftliche Untersuchung umfasst die Untersuchungsschritte Fragestellung, Hypothese, Planung und Durchführung einer Untersuchung sowie Datengewinnung und –auswertung. Es liegen 52 % der Testaufgaben im MC-Format, 7 % als Kurzantwort und 41 % als erweiterte Antwort vor. In Abbildung 17 wird eine Beispielaufgabe aus dem Kompetenzbereich Erkenntnisgewinnung dargestellt.

Fleißiges Lieschen

Naturwissenschaftlicher Einführungstext

Max will die Wachstumsbedingungen von Pflanzen untersuchen. Er verwendet als Versuchspflanze das „Fleißige Lieschen“, eine pflegeleichte und widerstandsfähige Zimmerpflanze.

Beispielaufgabe:

Er führt folgendes Experiment durch:

Faktoren	Pflanze 1	Pflanze 2
Temperatur	25 °C	25 °C
Ort	am hellen Fenster	im dunklen Schrank
gießen	1-mal täglich	1-mal täglich
düngen	2-mal wöchentlich	2-mal wöchentlich

Nenne die Hypothese (Vermutung), die dem Experiment von Max zugrunde liegt.

Abbildung 17: Beispielaufgabe zum Prozesswissen in der Biologie

Diese Aufgabe im erweiterten Antwortformat ist dem Themengebiet naturwissenschaftliche Untersuchung zuzuordnen. Die Schülerinnen und Schüler müssen nach dem Aufgabenentwicklungsmodell einen Zusammenhang organisieren. Die zu bildende Hypothese beinhaltet ein Verständnis des Zusammenhangs zwischen dem Wachsen der Pflanze und der Variable Licht.

Auch für den Kompetenzbereich Erkenntnisgewinnung wird in der folgenden Abbildung 18 die Verteilung der Schülerinnen und Schüler auf die Anzahl der bearbeiteten Testaufgabe dargestellt. Die mittlere Anzahl von Personen pro Aufgabe beträgt 254 Schülerinnen und Schüler ($\text{Min}_{\text{Anzahl Personen}} = 241$, $\text{Max}_{\text{Anzahl Personen}} = 488$, $\tilde{x}_{\text{Anzahl Personen}} = 250$).

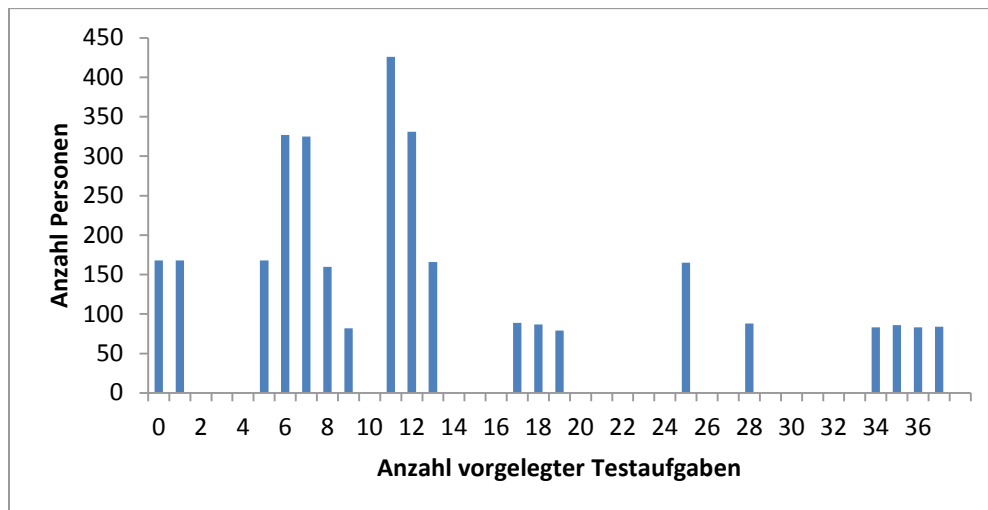


Abbildung 18: Anzahl vorgelegter Testaufgaben für den Kompetenzbereich Erkenntnisgewinnung

Von 3 165 Schülerinnen und Schülern haben 2 977 Testaufgaben zum Kompetenzbereich Erkenntnisgewinnung vorgelegt bekommen. Von diesen haben 24 % der Schülerinnen und Schüler elf oder zwölf Aufgaben vorgelegt bekommen, weitere 21 % bearbeiteten sechs bis sieben Testaufgaben. Die Spanne reicht bis zu 37 vorgelegten Testaufgaben. Die Schwierigkeitsverteilung der Testaufgaben für den Kompetenzbereich Erkenntnisgewinnung ist in Abbildung 19 – wiederum auf einer Logit-Skala in einem 1PL-Modell – dargestellt.

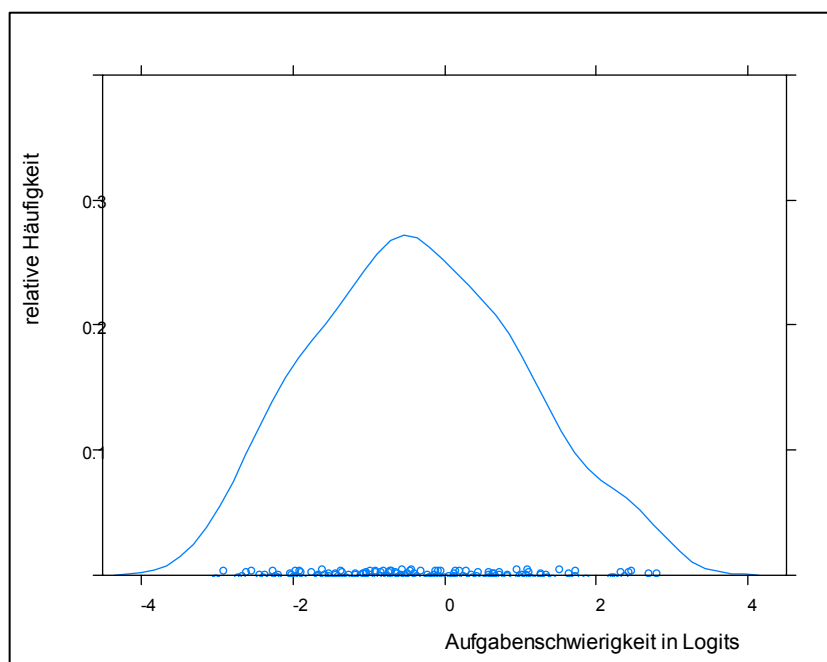


Abbildung 19: Schwierigkeitsverteilung der Testaufgaben zu Prozesswissen in der Jahrgangsstufe 10
Anmerkungen. Jeder Punkt steht für eine Testaufgabe.

Die mittlere Schwierigkeit der Aufgaben beträgt $M_b = -0.33$ ($SD_b = 1.34$, $Max_b = 2.80$, $Min_b = -3.01$). Die modellbasierte Reliabilität der Skala Biologie Prozesswissen liegt bei McDonalds $\omega = .97$.

Wie bereits erwähnt wurde, können die Schülerinnen und Schüler alle zur Lösung relevanten Informationen dem einleitenden biologischen Text entnehmen. Dieses Vorgehen bedeutet aber auch, dass ein nicht unbeträchtlicher Anteil zur Lösung der Aufgabe durch die sprachlichen Kompetenzen der Schülerinnen und Schüler bestimmt werden könnte. Auch die Tatsache, dass die Schülerinnen und Schüler mittels eines *paper-and-pencil* Tests getestet wurden, trägt zu dieser Problematik bei. Um die Zusammenhänge mit den sprachlichen Kompetenzen und weiteren relevanten Konstrukten zu testen, wurden Validierungsinstrumente eingesetzt, die im Folgenden vorgestellt werden sollen.

4.3.2 Instrumente zur Validierung der Kompetenzmessung

Die *sprachlichen Kompetenzen* werden aus testökonomischen Gründen mit Hilfe von zwei sogenannten Lückentests erhoben (Ackerman et al., 2000). Es war nicht möglich, die sprachlichen Kompetenzen in einem ähnlichen Umfang wie die naturwissenschaftliche Kompetenz zu testen. Der gewählte Lückentest ist im Vergleich zu weiteren Kurz-Tests wie z. B. dem Lesegeschwindigkeits- und -verständnistest (LGVT) (Schneider, Schlagmüller & Ennemoser, 2007) in besonderem Maße für die Testung der sprachlichen Kompetenzen geeignet. Lückentests bestehen aus kurzen Texten, die jeweils 20 Wortlücken aufweisen, die von den Schülerinnen und Schülern ausgefüllt werden müssen. Dabei fehlen gleich viele Buchstaben eines Wortes, wie gegeben sind, oder es fehlt ein Buchstabe mehr. So muss z. B. in einem Satzzusammenhang das unvollständige Wort *erkl_____* als erklären erkannt werden und die Lücke dementsprechend fehlerfrei ausgefüllt werden. Jeder Text wurde von den Schülerinnen und Schülern in jeweils vier Minuten bearbeitet. Die insgesamt 40 Testaufgaben wurden im Mittel von 2 979 Schülerinnen und Schülern beantwortet ($SD = 155$). Um zu vertretbaren Modellfitwerten für die Messmodelle zu gelangen, mussten von beiden Texten Variablen entfernt werden. Als Kriterium für die Eliminierung wurden Lösungsraten von über 95 % bzw. unter 5 % gesetzt. Für den zweiten Lückentest führte dieses Verfahren nicht zu vertretbaren Modellfitwerten. Daher wurde mit diesem Test eine *sequential explorative factor analysis* (SEFA) durchgeführt (Kano & Harada, 2000). Die Anzahl der entfernten Testaufgaben für den jeweiligen Test wird in Tabelle 18 berichtet. Des Weiteren werden für beide Texte die Werte zur Schwierigkeitsverteilung der Testaufgaben sowie die Modellfitwerte berichtet. Der Bentler CFI und der *Tucker-Lewis Index* (TLI) sind inkrementelle Fitwerte, mittels derer die relative Verbesserung der Passung eines spezifizierten Modells im Vergleich zu einem gedachten *baseline*-Modell berechnet werden, welches keinerlei

Kovarianzen zwischen den latenten Variablen annimmt (Bentler, 1990; Kline, 2011). TLI und CFI können Werte auf einem Kontinuum von Null bis Eins annehmen. Für eine gute Modellpassung sollte sie Werte von .95 und höher aufweisen. Der RMSEA ist ein sogenannter *badness-of-fit index*. Er sollte bei .05 und niedriger liegen. Ein RMSEA größer .10 deutet auf ernsthafte Probleme hin (Kline, 2011).

Tabelle 18: Skalenbeschreibung der Lückentests – entfernte Aufgaben, Schwierigkeitsverteilung der Testaufgaben sowie Modellfitwerte der beiden Messmodelle

	Lückentest 1	Lückentest 2
Anzahl entfernter Testaufgaben	6	6
Schwierigkeitsverteilung		
<i>M</i>	-0.72	-0.69
<i>SD</i>	0.50	0.66
<i>Mdn</i>	-0.71	-0.69
Min	-1.72	-1.60
Max	0.34	0.59
Modellfitwerte		
RMSEA	.04	.06
Bentler CFI	.93	.95
TLI	.92	.94

Anmerkungen. *M* = arithmetisches Mittel, *SD* = Standardabweichung, *Mdn* = Median, Min = Minimum, Max = Maximum, RMSEA = *Root Mean Square of Approximation*, Bentler CFI = Bentler *Comparative Fit Index*, TLI = *Tucker-Lewis Index*, WRMR = *Weighted Root Mean Square Residual*.

Von den 20 Variablen des ersten Lückentests mussten drei Variablen aufgrund eines Fehlers in der Kodierung eliminiert werden. Weitere drei Variablen zeigten bezogen auf die gesetzten Kriterien ungenügende Testwerte. Im zweiten Lückentest lag eine Variable mit Kodierungsfehler und weitere fünf Variablen wurden aus dem Test entfernt. Der CFL, TLI und RMSEA bestätigen einen guten Modellfit. Für beide Lückentests liegt der TLI knapp außerhalb des optimalen Wertebereichs. Der RMSEA verfehlt für den Lückentest 2 die .05 knapp. Die beiden finalen Lückentests bestehen aus 28 Aufgaben und weisen eine Reliabilität von McDonalds $\omega_{LT1} = .90$ und McDonalds $\omega_{LT2} = .91$ auf.

Zur Erfassung der *kognitiven Grundfähigkeit* wurde der häufig eingesetzte KFT verwendet (Heller & Perleth, 2000). Um sicherzustellen, dass die Erhebung der kognitiven Grundfähigkeit von den sprachlichen Kompetenzen der Schülerinnen und Schüler abgegrenzt werden kann, wurde der non-verbale N2-Untertest verwendet. Da die Studie zu Beginn der Jahrgangsstufe 10 durchgeführt wurde, wurden die 30 Testaufgaben für die 9. und 10. Klassenstufe verwendet. Die Testpersonen bekamen

pro Testaufgabe ein geometrisches Figurenpaar vorgelegt, z. B. einen kleinen weißen Kreis und einen großen schwarzen Kreis. Daraufhin wurde eine weitere Figur, z. B. ein kleines weißes Rechteck vorgegeben. Die Schülerinnen und Schüler mussten dann aus fünf weiteren Figuren diejenige auswählen, die sich zu der gegebenen Figur verhält wie die zweite Figur zur ersten des gegebenen Figurenpaares. Zur Bearbeitung dieser Testaufgaben standen den Schülerinnen und Schülern 10 Minuten zur Verfügung. Im Mittel haben 3 045 Schülerinnen und Schüler die 30 Testaufgaben beantwortet ($SD = 93$). Die mittlere Schwierigkeit der KFT-Aufgaben liegt bei $M_b = -0.50$ ($Mdn_b = -0.51$, $SD_b = 0.43$, $Min_b = -1.37$, $Max_b = 0.41$). Die Modellfitwerte $RMSEA = .03$, Bentler CFI = .96 und TLI = .96 verdeutlichen die gute Passung des Messmodells. Die Reliabilität der KFT-Skala liegt bei McDonalds $\omega = .93$.

Die *schulischen Leistungen der Schülerinnen und Schüler* wurden über die von den Schulleitungen bzw. von den Schulkoordinatorinnen und -koordinatoren berichteten Fachnoten eingeholt. Mittels einer Listenabfrage wurde erfasst, welche Note die einzelne Schülerin bzw. der einzelne Schüler auf dem Zeugnis am Ende des letzten Schuljahres (Schuljahr 2008/09) in den Fächern Biologie, Chemie, Physik, Mathematik und Deutsch hatte. Falls die Schülerin bzw. der Schüler in diesem Schuljahr keinen Unterricht in einem Fach erhalten hat, liegen keine Informationen vor. Das Schulpersonal musste die Fachnote notieren oder ein Kästchen ankreuzen, falls die betreffende Schülerin bzw. der betreffende Schüler in diesem Fach im fraglichen Zeitraum keinen Unterricht erhalten hatte oder nicht benotet wurde. Da in den einzelnen Bundesländern und Schulformen der Benotung unterschiedlich variierende Skalen zu Grunde liegen, wurden die Angaben zu den Schulnoten auf eine einheitliche Skala von eins bis sechs übertragen. Um die Interpretation der Ergebnisse zu erleichtern, wurden die Schulnoten zusätzlich in umgekehrter Reihung kodiert, sodass höhere Werte besseren Fachnoten entsprechen und positive Zusammenhänge der Fachnoten mit anderen Leistungsmaßen als solche interpretiert werden können. Des Weiteren wird davon ausgegangen, dass der relationale Bezugsrahmen zur Benotung innerhalb einer Klasse angesetzt wird und die Lehrkräfte zwischen den Klassen nicht zwangsläufig denselben Benotungsmaßstab ansetzen und daher mit unterschiedlicher Strenge zensieren können. Daher wurden die rekodierten und umgepolten Noten zusätzlich in der Gesamtstichprobe am Klassenmittelwert zentriert. Da das gleiche Verfahren nicht für die Leistungswerte angewendet wurde, könnten die Zusammenhänge zwischen Fachnoten und Leistungswerten systematisch unterschätzt werden. Die Anzahl der Schülerinnen und Schüler, zu denen Informationen zu den unterschiedlichen Fachnoten vorliegen, kann Tabelle 19 entnommen werden. Des Weiteren werden in Tabelle 19 auch die Anzahl der Schülerinnen und Schüler berichtet, die im getesteten Klassenverband unterrichtet wurden.

Tabelle 19: Anzahl der Schülerinnen und Schüler mit Angaben zu den Schulnoten und zum Unterricht im Klassenverband pro akademische Disziplin (absolut, Klammern in Prozent)

Akademische Disziplin	<i>n</i>	Klassenverband
Biologie	2 592	2 535 (98)
Chemie	2 800	2 537 (91)
Physik	2 527	2 521 (100)
Mathematik	3 123	2 769 (89)
Deutsch	3 109	k. A.

Anmerkungen. k. A. = keine Angaben.

Schülerinnen und Schüler, die Biologietesthefte vorgelegt bekommen haben (s. Seite 104), haben auch einen Schülerfragebogen mit Fragen zu ihrem *akademischen Selbstkonzept* in Biologie und einer der beiden weiteren naturwissenschaftlichen Disziplinen vorgelegt bekommen. Daher liegen für das Fach Biologie höhere Antwortraten vor als in den beiden anderen Fächern. Dieser Fragebogen war in zwei Bundesländern (Berlin und Thüringen) für die Schülerinnen und Schüler verpflichtend, in den verbleibenden sechs freiwillig.

Die Skalen zum akademischen Selbstkonzept wurden dem in der PISA-Erhebung 2006 verwendeten Fragebogen entnommen und an die akademischen Disziplinen Biologie, Physik und Chemie angepasst. Die Schülerinnen und Schüler mussten Aussagen wie z. B. „Ich glaube, dass ich anspruchsvollen Stoff in Biologie leicht lernen kann“ zustimmen oder ablehnen. Die gleichen Aussagen wurden auch für Physik und Chemie abgeändert. Die Skala zum Selbstkonzept besteht pro naturwissenschaftlicher Disziplin aus sechs Testitems, die auf einer vierstufigen Likert-Skala mit den Ausprägungen „stimme ganz zu“ bis „stimme gar nicht zu“ beantwortet werden konnten. Die Messmodelle der Selbstkonzeptskalen wurden in getrennten eindimensionalen 2PL-Modellen berechnet (s. Seite 122). Die Selbstkonzeptskala wurde in Biologie von 2 627 Schülerinnen und Schülern beantwortet und weist eine Reliabilität von McDonalds $\omega = .95$ auf. Die standardisierten Faktorladungen des eindimensionalen 2PL-Modell liegen im Mittel bei $\lambda_{\text{Mittel}} = .87$ ($\lambda_{\text{SD}} = 0.03$). Die Physik-Skala wurde von 1 378 Schülerinnen und Schülern beantwortet und die Faktorladungen liegen im Mittel bei $\lambda_{\text{Mittel}} = .90$ ($\lambda_{\text{SD}} = 0.02$). In Chemie liegen Antworten von 1 513 Schülerinnen und Schülern vor. Die gemittelten Faktorladungen liegen bei $\lambda_{\text{Mittel}} = .90$ ($\lambda_{\text{SD}} = 0.02$). Die Reliabilitäten der Chemie- und Physik-Selbstkonzeptskalen liegen jeweils bei McDonalds $\omega = .96$.

4.4 Statistische Analysen

Den formulierten Fragestellungen liegt das Kompetenzkonzept zu Grunde. Kompetenzen sind nicht direkt beobachtbare – so genannte latente – Merkmale von Personen. Diese können nur über das Erfassen von beobachtbaren Merkmalen als Indikatoren für das latente Merkmal sichtbar gemacht werden (Köller, 2008). Da die latenten Variablen nicht direkt sichtbar sind, müssen diese über geeignete Indikatoren in einem Messmodell operationalisiert werden. Die verwendete Analysemethode zur Beantwortung der Forschungshypothesen und zur Spezifikation der Messmodelle – das Modellieren der Kompetenzen über strukturprüfende Verfahren auf latenter Ebene – ermöglicht es, über manifeste Indikatoren diese latenten Faktoren zu modellieren. Die latenten Faktoren spiegeln hierbei das zu erfassende, nicht beobachtbare Merkmal wider. Die Messfehlereinflüsse bei der Messung dieses Merkmals werden berücksichtigt, indem ein manifester Testwert in einen „wahren“ Wert und einen Messfehlerwert zerlegt wird (Eid, 1999). Theoretische Vorüberlegungen zu Beziehungen zwischen den latenten Variablen ermöglichen Analysen, die in den strukturprüfenden Verfahren spezifiziert werden können (Backhaus, 2006). Auf Grundlage der Indikatorvariablen der latenten Variablen werden Kovarianzen und Korrelationen zwischen den Indikatorvariablen verwendet, welche zur Bestimmung der Beziehungen zwischen den latenten und den Indikatorvariablen und zwischen den latenten Variablen herangezogen werden. Alle Messmodelle zur Beantwortung der Fragestellungen werden mit der Software Mplus 6.1 (Muthén & Muthén, 1998-2010) berechnet. Hierbei werden jeweils reflexive Messmodelle zu Grunde gelegt (Edwards & Bagozzi, 2000).

4.4.1 Die Modellierung der Messmodelle

Das Skalenniveau der manifesten Indikatoren ist nicht einheitlich; d. h. einige Konstrukte werden durch intervallskalierte bzw. näherungsweise intervallskalierte Indikatoren beschrieben, andere hingegen durch Indikatoren mit dichotomen Skalenniveau. Dieser Unterschied wird im Messmodell berücksichtigt. Werden Konstrukte, wie in der vorliegenden Arbeit das Selbstkonzept in den drei naturwissenschaftlichen Disziplinen, durch intervallskalierte bzw. näherungsweise intervallskalierte Indikatoren beschrieben, wird ein essentiell tau-äquivalentes oder tau-kongenerisches Modell verwendet (Kline, 2011). Werden Konstrukte, wie in der vorliegenden Arbeit das Konzept- und Prozesswissen sowie die sprachlichen Kompetenzen und die kognitive Grundfähigkeit, durch dichotome Indikatoren beschrieben, wird statt dessen ein aus der *Item Response Theory* (IRT) abgeleitetes 1PL- oder 2PL-Modell verwendet. Hierbei werden aus den dichotomen Antwortmustern Antwortwahrscheinlichkeiten aus den relativen Antworthäufigkeiten geschätzt und diese dann in logarithmierte Wettquotienten transformiert. Die so gewonnene Skala wird als Logit-Skala

bezeichnet, ihre Werte als Logits (Rost, 2004). Auf dieser Skala können sowohl Schwierigkeiten von Testaufgaben als auch Fähigkeitswerte von Personen abgetragen werden, wodurch beide Parameter direkt zueinander in Beziehung gesetzt werden können (Embretson & Reise, 2000)

In den berichteten Messmodellen (s. Seite 115) werden sowohl für die dichotomen als auch für die (näherungsweise) intervallskalierten Indikatoren die Faktorladungen entweder auf einen gemeinsamen Wert fixiert oder frei geschätzt. Inhaltlich bedeutet diese Unterscheidung entweder, dass die Indikatoren zu gleichen Teilen, oder unterschiedlich stark gewichtet zur Varianzaufklärung des Faktors beitragen. Um empirisch zu prüfen, welche der beiden Varianten eine bessere Passung auf die vorliegenden Daten erzielt, können probeweise beide Möglichkeiten modelliert und anschließend das *Bayesian Information Criterion* (BIC) beider Varianten verglichen werden. In Tabelle 20 werden die BIC der jeweils konkurrierenden Modelle gegenübergestellt. Ein Modell mit einem geringeren BIC als ein weiteres, hierzu genestetes Modell passt besser auf die Daten und ist daher zu favorisieren.

Tabelle 20: BIC der 1PL- und 2PL-Modelle bzw. der essentiell tau-äquivalenten und tau-kongenerischen Modelle für die der Untersuchung zu Grunde liegenden Konstrukte Konzept- und Prozesswissen, sprachliche Kompetenzen, kognitive Grundfähigkeit und akademisches Selbstkonzept

Konstrukt		1PL-Modell	2PL-Modell
Konzeptwissen		30 212.62	30 905.99
Prozesswissen		43 139.02	43 971.50
Lesetest 1		40 641.69	40 393.33
Lesetest 2		37 938.82	37 023.89
Kognitive Grundfähigkeit		96 540.95	95 940.52

	Disziplin	Essentiell tau-äquivalentes Modell	Tau-kongenerisches Modell
Selbstkonzept	Bio	26 810.54	26 783.09
	Phy	14 593.32	14 572.34
	Che	16 190.73	16 169.68

Anmerkungen. 1PL = *one parameter logistic*, 2PL = *two parameter logistic*.

Die 1PL-Modelle weisen für die Testaufgaben einen niedrigeren BIC auf. Die 2PL-Modelle weisen für die Validierungsinstrumente einen niedrigeren BIC auf als die 1PL-Modelle. Daher werden für die Testaufgaben 1PL-Modelle verwendet. Die Faktorladungen der Testaufgaben auf den latenten Faktor werden also festgesetzt. Diese Modelle entsprechen den aus der IRT-Nomenklatur bekannten

Raschmodellen. Die in Kapitel 4.3.1.2 beschriebenen Aufgabenschwierigkeiten beziehen sich auf diese Raschmodelle. Für die Validierungsinstrumente weisen die BIC der 2PL-Modelle niedrigere Werte auf. Aus diesem Grund werden tau-kongenerische oder 2-PL-Modelle für die Validierungsinstrumente zu Grunde gelegt. Die tatsächlich verwendeten Messmodelle werden im Folgenden näher beschrieben und können zudem den Abbildungen im Anhang (Abbildung 26 bis Abbildung 33) entnommen werden.

Die beiden Messmodelle für Konzept- und Prozesswissen werden durch 117 bzw. 153 dichotome manifeste Indikatoren beschrieben. Im verwendeten eindimensionalen 1PL-Modell werden alle Faktorenladungen auf einen gemeinsamen Wert geschätzt. Dies bedeutet, dass jeder Indikator mit demselben Gewicht zur Varianzaufklärung des Faktors beiträgt. Die Varianz des latenten Faktors wird zur Normierung der Metrik auf Eins gesetzt. Die Daten der Schülerinnen und Schüler liegen dichotom vor. Die berichteten Schwierigkeiten der Testaufgaben (s. Seite 115) entsprechen also dem Punkt auf der Logit-Skala, für den die Lösungswahrscheinlichkeit der Testaufgabe 50 % beträgt.

Die Messmodelle der Validierungsinstrumente unterscheiden sich voneinander. Die beiden Messmodelle für sprachliche Kompetenzen und kognitive Grundfähigkeit werden durch 28 bzw. 30 dichotome manifeste Indikatoren abgebildet. Auch für diese beiden Instrumente liegen die Daten der Schülerinnen und Schüler dichotom vor. Die eindimensionalen Modelle wurden analog zu den Modellen zu den Testaufgaben spezifiziert. Der einzige Unterschied ist, dass die Faktorladungen frei geschätzt werden; sie sind somit 2PL-Modelle. Die akademischen Selbstkonzepte werden jeweils durch sechs manifeste (näherungsweise) intervallskalierte Indikatoren gemessen. Auch in diesen Modellen werden die Faktorladungen frei geschätzt und die Varianz des latenten Faktors auf Eins gesetzt. Sie werden daher als tau-kongenerische Modelle spezifiziert.

4.4.2 Die Modellierung der Strukturmodelle

In den Strukturmodellen werden die latenten Faktoren der Messmodelle entsprechend der aufgestellten Forschungshypothesen in Zusammenhang gebracht. Dabei können sowohl gerichtete als auch ungerichtete Zusammenhänge modelliert werden. Die vorliegenden Daten bringen einige Herausforderungen mit sich. Wie diesen begegnet wurde, wird im Folgenden beschrieben.

Die hohen Intraklassenkorrelationen der Testleistungen in Biologie auf Ebene der Klasse und der Schule müssen in den Strukturmodellen berücksichtigt werden. Für die Fragestellungen sind die Klassenkompositionseffekte und Effekte auf Schulebene nicht relevant. Da die Varianz von Schülerinnen und Schülern gleicher Klassen kleiner ist als die Varianz von Schülerinnen und Schülern verschiedener Klassen, werden jedoch die Standardfehler für Parameter auf der Schülerebene

systematisch unterschätzt (Luke, 2005; Wolter, 1985). Mplus bietet die Möglichkeit, mit der *Cluster*-Funktion diese Unterschätzung der Standardfehler zu korrigieren. Diese Funktion wurde in allen Modellierungen (Messmodelle und Fragestellungen) auf Ebene der Klasse angewendet. Mplus bietet nur die Möglichkeit der Beachtung einer Clusterebene. Die Klassenebene wurde gewählt, da diese die niedrigste Ebene darstellt, auf der gruppenspezifische Homogenität innerhalb der Stichprobe vorliegen kann. Sie fängt damit einen Teil der Homogenitäten auf der nächsthöheren Ebene – der Schulebene – ab. Für das Modell mit den klassenzentrierten Fachnoten wurde die Schule gewählt. Da die Fachnoten bereits am Klassenmittelwert zentriert wurden, wird für Modelle, in denen die Fachnoten berücksichtigt werden, die Schule als *Cluster*-Ebene verwendet. In einem Modell, in das die Schulnoten der Schülerinnen und Schüler einfließen, wird mit der Zentrierung auf Klassenebene eine geringfügige Unschärfe in Kauf genommen. In Biologie wurden z. B. 2 % der Schülerinnen und Schüler nicht im getesteten Klassenverband unterrichtet. Die Zentrierung wird als auch für die 2 % Schülerinnen und Schüler vorgenommen, die nicht im Klassenverband getestet wurden.

Die Hinzunahme von Fachnoten in die Strukturmodelle legt auf den ersten Blick die Überlegung der Verwendung von Mehrebenenmodellen nahe. Dies ist jedoch für die vorliegenden Daten nicht zielführend. Zunächst werden nicht alle Schülerinnen und Schüler im Klassenverband in den Naturwissenschaften unterrichtet. Die Schülerinnen und Schüler werden an einem bestimmten Tag in einer bestimmten Klassenkonstellation getestet. In der Konstellation der Testung könnte eine Klasse in dem gegebenen Schuljahr z. B. zusammen Physik unterrichtet bekommen, eine weitere Klasse könnte Biologie in der getesteten Konstellation unterrichtet bekommen. Je nach Klasse liegen den Daten somit unterschiedliche Konstellationskompositionen zu Grunde, die nicht in einer Mehrebenenanalyse abgebildet werden können. Wenn die Schülerinnen und Schüler nun Fachnoten zu mehreren naturwissenschaftlichen und nicht-naturwissenschaftlichen Fächern angeben, erhalten sie diese in unterschiedlichen Klassenverbänden und von unterschiedlichen Lehrkräften. In Biologie und Chemie tritt dieses Problem stärker auf als in Physik. Diese methodischen Probleme und die bereits im theoretischen Teil der Arbeit beschriebenen theoretischen Probleme führten zu der Entscheidung, keine Mehrebenenmodelle zu verwenden. Diese Modelle ermöglichen es nicht, die benannten methodischen Probleme zu beheben.

Eine weitere Herausforderung der Daten ergibt sich aus dem PBIBD. Dieses Design führt zwangsläufig zu fehlenden Werten, mit denen bei der Berechnung der Mess- und Strukturmodelle umgegangen werden muss. Die Schülerinnen und Schüler haben die Testhefte randomisiert zugewiesen bekommen, sodass die aus dem Testdesign resultierenden fehlenden Werte als *missing completely at random* (MCAR) vorliegen. Für diese unsystematischen Ausfallprozesse bietet Mplus den *full information maximum likelihood* Schätzer (FIML). Bei dieser Schätzmethode werden die

Parameter und deren Standardfehler nur auf Basis der beobachteten Werte vorgenommen und keine fehlenden Werte imputiert oder geschätzt (Lüdtke, Robitzsch, Trautwein & Köller, 2007). Es wurden keine Hilfsvariablen bei der Schätzung der Parameter berücksichtigt.

Die manifesten Indikatoren der latenten Konstrukte liegen dichotom vor, im Fall der Fachnoten (näherungsweise) intervallskaliert. Für die Schätzung der Mess- und Strukturmodelle wurde der MLR verwendet, um den beiden Bedingungen (a) keine Normalverteilung und (b) Multi-Matrix Design gerecht werden zu können. Der MLR-Schätzer ist robust gegen eine Verletzung der Normalverteilung auf manifester Ebene (Kline, 2011).

Für die Konstrukte Konzept- und Prozesswissen in Biologie, sprachliche Kompetenzen, kognitive Grundfähigkeit sowie für die Selbstkonzeptskalen wurden jeweils einzelne Messmodelle berechnet. Für die Selbstkonzeptskalen musste zunächst geklärt werden, ob diese eine fachspezifische Ausprägung vorweisen oder ob sie nur für eine globale Naturwissenschaft zu bestimmen sind. Hierzu wurde das BIC eines eindimensionalen Modells des Selbstkonzepts in den Naturwissenschaften jeweils dem BIC eines dreidimensionalen Modells der Selbstkonzepte in Biologie, Chemie und Physik gegenübergestellt (s. Tabelle 21). Für den relativen Modellfit-Wert BIC liegt – im Gegensatz zu den absoluten Modellfitwerten wie z. B. dem TLI – kein absoluter Wert vor, der die Güte eines Modells sicherstellen könnte. Vielmehr zeigt erst der Vergleich von BIC-Werten unterschiedlicher Modelle, welches am besten den erhobenen Daten entspricht. Der BIC wurde gewählt, da dieser die Anzahl der Parameter eines Modells in die Evaluation des Modellfits mit einbezieht.

Tabelle 21: Gegenüberstellung der BIC der ein- und dreidimensionalen Analysen für die naturwissenschaftlichen Selbstkonzepte

Konstrukt	Modell	BIC
Selbstkonzept	Eindimensional	65 817.42
	Dreidimensional	57 144.94

Anmerkungen. BIC = *Bayesian Information Criterion*.

Für die naturwissenschaftlichen Selbstkonzepte liegt Dreidimensionalität vor. Die Ladungen auf den drei Selbstkonzeptfaktoren liegen zwischen $\lambda = .83$ und $\lambda = .93$. Die Korrelation zwischen den Faktoren Selbstkonzept Biologie und Chemie beläuft sich auf $r = .46$ und zwischen Selbstkonzept Chemie und Physik auf $r = .45$. Die Faktoren Selbstkonzept Biologie und Physik hängen mit $r = .35$ zusammen. Diese Korrelationen liegen nah an den vorgefundenen Zusammenhängen bei Jansen et al. (in Revision).

5 Ergebnisse

Nach einer kurzen Beschreibung der Stichprobe werden die Ergebnisse chronologisch entlang der Forschungshypothesen berichtet. Die Stichprobe besteht zu 51 % Mädchen und das durchschnittliche Alter beträgt 15.50 Jahre ($SD = 2.16$). Zu den sprachlichen Kompetenzen beantworteten die Schülerinnen und Schüler durchschnittlich 30 von 40 Testaufgaben ($SD = 7$), in der kognitiven Grundfähigkeit 20 von 30 Testaufgaben ($SD = 6$). Sie lösten damit mehr als die Hälfte der Tests. Das mittlere Selbstkonzept der Schülerinnen und Schüler ist in der Biologie mit $M = 2.86$ ($SD = 0.67$) im Vergleich zu den anderen beiden akademischen Disziplinen am höchsten. Die Selbstkonzepte in Chemie nehmen mit $M = 2.51$ ($SD = 0.77$) und in Physik mit $M = 2.55$ ($SD = 0.79$) geringfügig niedrigere Werte an. Die Fachnoten der Schülerinnen und Schüler werden in Tabelle 22 aufgeführt.

Tabelle 22: Durchschnittliche Fachnote der Schülerinnen und Schüler in den naturwissenschaftlichen Disziplinen sowie in Deutsch und Mathematik

Fach	M	SD
Deutsch	2.93	0.82
Mathematik	3.08	0.95
Biologie	2.83	0.88
Chemie	2.95	0.94
Physik	2.93	0.90

Anmerkungen. Es wurden die rekodierten, nicht zentrierten und nicht umgepolten Fachnoten verwendet. Niedrigere Werte stehen somit für bessere Leistungen.

M = arithmetisches Mittel; SD = Standardabweichung.

Bei den Fachnoten hat die Biologie den höchsten Wert zu verzeichnen. Im Fach Mathematik haben die Schülerinnen und Schüler den geringsten Notendurchschnitt. Dazwischen liegen die weiteren naturwissenschaftlichen Fächer und Deutsch. Die Leistungen in den naturwissenschaftlichen Fächern liegen nah beieinander.

Tabelle 23 gibt einen Überblick über die Korrelationen zwischen den Summenwerten (für Lückentest und KFT) sowie den Mittelwerten (für Selbstkonzept und Fachnoten) der eingesetzten Validierungsinstrumente.

Tabelle 23: Produkt-Moment-Interkorrelationen der Summenwerte des Lückentests, KFT, Selbstkonzepts und der Fachnoten (Validierungsinstrumente)

	Fach	Lückentest	KFT	Selbstkonzept			Noten			
				Bio	Che	Phy	Bio	Che	Phy	Ma
Lückentest										
KFT		.42*								
Selbstkonzept	Bio	.15*	.15*							
	Che	.00	.14*	.42*						
	Phy	-.03	.15*	.32*	.35*					
Fachnote	Bio	.12*	.12*	.33*	.22*	.14*				
	Che	.10*	.16*	.21*	.43*	.22*	.56*			
	Phy	.09*	.16*	.19*	.28*	.37*	.51*	.59*		
	Mat	.10*	.26*	.31*	.06*	.08*	.45*	.57*	.59*	
	Deu	.21*	.09*	.16*	.06	.01	.51*	.42*	.43*	.39*

Anmerkungen. * = Die Korrelation ist auf einem Niveau von $\alpha = .01$ (2-seitig) statistisch signifikant von Null verschieden.

KFT = Kognitiver Fähigkeitstest.

Die sprachlichen Kompetenzen korrelieren erwartungsgemäß am höchsten mit den schulischen Leistungen in Deutsch. Innerhalb der naturwissenschaftlichen Disziplinen liegt der Zusammenhang für die Biologie am höchsten. Die kognitive Grundfähigkeit hängt bezogen auf die schulischen Leistungen am höchsten mit den Leistungen in Mathematik zusammen, bezogen auf die naturwissenschaftlichen Disziplinen am geringsten mit der Biologie. Die Selbstkonzepte einer akademischen Disziplin hängen zudem höher mit den schulischen Leistungen in dieser als mit den anderen akademischen Disziplinen zusammen. Dieser Befund deckt sich mit den internen Vergleichsprozessen des I/E-Modells (z. B. Marsh & Köller, 2004). Eine Ausnahme bildet hier das Selbstkonzept in Biologie, das mit den schulischen Leistungen in Mathematik ähnlich korreliert wie mit den Leistungen in Biologie. Die schulischen Leistungen untereinander korrelieren dagegen unabhängig von den akademischen Disziplinen hoch miteinander. Besonders hohe Zusammenhänge zeigen sich zwischen Biologie und Chemie ($r = .56$), Chemie und Physik ($r = .59$) sowie Physik und Mathematik ($r = .59$). Auch dieser Befund deckt sich mit bereits berichteten Zusammenhangsmustern zwischen Schulnoten (Möller et al., 2007). Die Ergebnisse zu den Hypothesen werden im Folgenden getrennt für die fachdidaktisch-kognitionspsychologische und der differentialpsychologische Perspektive berichtet.

5.1 Hypothesentestung

Die Hypothesentestung wird in der Reihenfolge der formulierten Fragestellungen vorgenommen. Da die Grundlage für alle Strukturmodelle eine hohe Anzahl an manifesten, dichotomen Variablen bzw. manifesten mehrkategorialen Variablen (Selbstkonzept) ist, berechnet Mplus keine χ^2 -Statistiken für die Modelle. Daher können diese sowie der Bentler CFI und der TLI für die Modelle nicht angegeben werden. Zunächst wird aus kognitionspsychologischer Sichtweise die Dimensionalität von Kompetenz in Biologie untersucht. Es folgen die differentialpsychologischen Fragestellungen.

5.1.1 Forschungshypothese A: Zusammenhang des Konzept- und Prozesswissens in Biologie – Fachdidaktisch-kognitionspsychologische Perspektive

Mit der Forschungshypothese A wird angenommen, dass Konzept- und Prozesswissen zwei empirisch trennbare, hoch korrelierende Teildimensionen von Kompetenz in Biologie sind. Um diese Hypothese zu testen, wurden die folgenden drei 1PL-Modelle mit Mplus spezifiziert und miteinander verglichen (s. Abbildung 20).

Modell Aa: *Globale Kompetenz in Biologie*. In diesem Modell laden alle 270 Aufgaben zu Konzept- und Prozesswissen in der Biologie auf einem latenten Faktor. Es wird eine eindimensionale Struktur von Kompetenz in Biologie angenommen.

Modelle Ab: Standardmodell – *Konzept- und Prozesswissen in Biologie*. Mit diesem Modell wird eine zwei-dimensionale Struktur mit den aus der Theorie hergeleiteten Dimensionen Konzept- und Prozesswissen angenommen. Die 117 Aufgaben zu Konzeptwissen laden auf einer latenten Konzeptwissensdimension, die 153 Aufgaben zu Prozesswissen laden auf einer zweiten Prozesswissensdimension.

Modell Ac: Zufallsmodell – *Zwei Zufallsfaktoren von Kompetenz in Biologie*. Dieses Modell dient der Absicherung, dass Modell Ab gegenüber dem Modell Aa nicht nur deswegen bessere Modellfit-Werte aufweist, weil es eine differenziertere Struktur als Modell Aa aufweist. Es werden wiederum zwei Faktoren modelliert. Die Aufgaben zu Kompetenz in Biologie werden jedoch zufällig auf die beiden Dimensionen aufgeteilt. Dieses Vorgehen führt zu einem Modell, in dem 117 Aufgaben auf der ersten Zufallsdimension und 153 Aufgaben auf einer zweiten Zufallsdimension laden.

Die Faktorladungen der Modelle werden festgesetzt (s. Tabelle 20). Im eindimensionalen Modell werden die Faktorladungen der manifesten Variablen des Konstrukts Konzeptwissen auf einen Wert festgesetzt, für die Faktorladungen der manifesten Variablen des Konstrukts Prozesswissen auf einen weiteren Wert. In den beiden zweidimensionalen Modellen werden die Faktorladungen auf einen Wert pro Dimension festgesetzt. Zur Normierung werden die Varianzen der latenten Faktoren auf Eins gesetzt; daher sind das Modell Aa und Ac nicht ineinander genestet. Für den Vergleich dieser beiden Modelle können daher nicht alle Modelltests verwendet werden. Die entsprechenden Modelle sind in Abbildung 20 graphisch dargestellt.

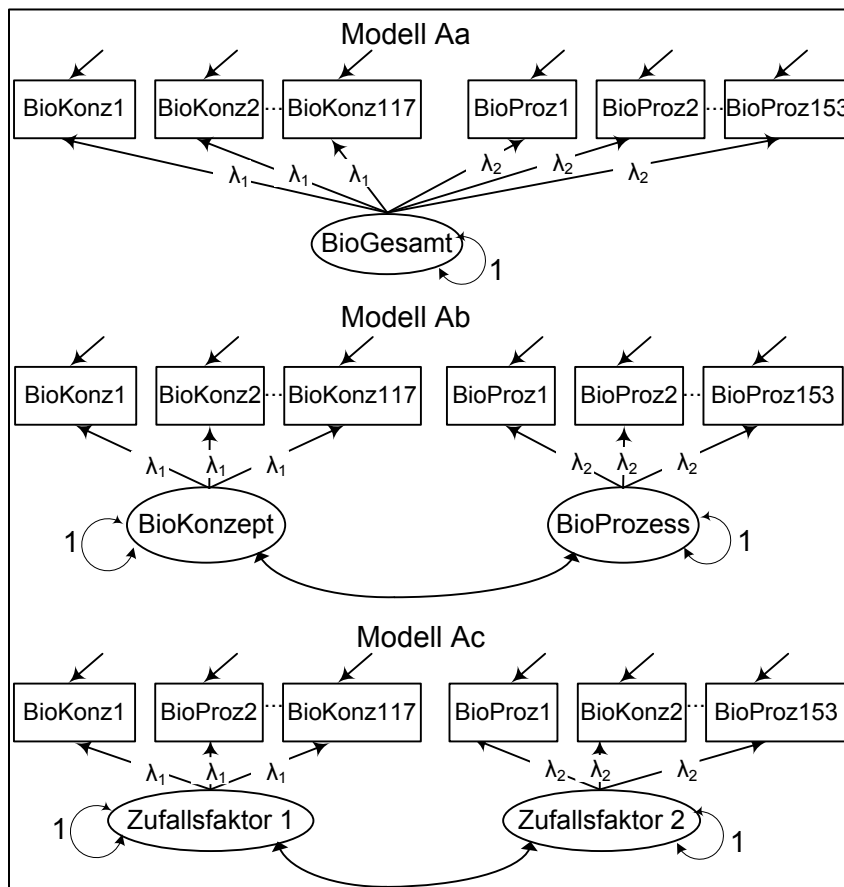


Abbildung 20: Spezifikation der Strukturmodelle Globale Kompetenz in Biologie (Modell Aa), Konzept- und Prozesswissen in Biologie (Ab) und zwei Zufallsdimensionen von Kompetenz in Biologie (Ac)

Anmerkungen. Bio = Biologie, Konz = Konzeptwissen, Proz = Prozesswissen.

Zum Vergleich der verschiedenen spezifizierten Modelle werden die unterschiedlichen Modellfit-Werte *loglikelihood* bzw. *final deviance*, *Akaike Information Criterion* (AIC) und BIC herangezogen. Auch für diese relativen Modellfit-Werte liegen keine absoluten Werte vor, die die Güte eines Modells sicherstellen könnten, sodass sie nur im Vergleich unterschiedlicher Modelle bewertet

werden können. Die Schätzung der Modellparameter erfolgt über eine *likelihood*-Funktion, die die Wahrscheinlichkeit der Daten bei gegebenen Modellparametern beschreibt. In einem iterativen Schätzverfahren werden die Modellparameter gesucht, für die der Funktionswert der *likelihood*-Funktion den höchsten Wert annimmt. Die *likelihood* ist der Logarithmus dieses maximalen Funktionswerts. Die *final deviance* als Maß für die Abweichung der tatsächlichen Daten von der Vorhersage des Modells ist $-2 \log(L)$ (Wilson & De Boeck, 2004). Bei dem Vergleich zweier Modelle, ist das Modell mit der geringeren *final deviance* zu favorisieren. Als zwei weitere Indikatoren für die Güte der drei Modelle werden das AIC und das BIC herangezogen. AIC und BIC ist jeweils eine um die Anzahl der Modellparameter adjustierte *final deviance* (Wilson & De Boeck, 2004). Das AIC ist hierbei im Vergleich zum BIC stärker von der Anzahl der Parameter eines Modells abhängig. Bei einem Vergleich von zwei Modellen spricht ein geringeres AIC und BIC für eine bessere Passung des Modells auf die Daten.

Zur weiteren Absicherung der Ergebnisse werden zwei Modelltests – der Wald-Test (Bollen, 1989) und der Satorra-Bentler χ^2 -Differenzentest (Satorra & Bentler, 2001) angewandt. Der Wald-Test wurde gewählt, da mit diesem Test einzelne Restriktionen eines Modells getestet werden können. Es kann z. B. eine vorgefundene hohe Korrelation zwischen zwei latenten Faktoren gegen eine perfekte Korrelation getestet werden. Der Satorra-Bentler χ^2 -Differenzentest setzt die Veränderung der *loglikelihood* und der geschätzten Parameter ins Verhältnis und prüft auf Grundlage eines χ^2 -Tests, ob diese Veränderung eine statistisch signifikante Verbesserung bedeutet. Bedingung zur Anwendung des χ^2 -Test zum Vergleich von Modellen sind kontinuierliche manifeste Variablen. Die vorliegenden Daten sind jedoch dichotom. Für diesen Fall bieten Satorra und Bentler (2001) eine Korrektur des χ^2 -Differenzentests an. Wenn der Signifikanzwert des Vergleichs eines liberalen Modells (im vorliegenden Fall das zweidimensionale Modell) mit einem restriktiveren Modell (im vorliegenden Fall das eindimensionale Modell) unter dem Signifikanzniveau von $\alpha = .05$ liegt, kann dieses Modell angenommen und das restriktivere Modell verworfen werden. Da Mplus aufgrund der hohen Anzahl an manifesten Variablen bei der Berechnung der Modelle keinen χ^2 -Test durchführt, beruht der χ^2 -Differenzentest auf den *loglikelihoods*. Das Modell, das im Vergleich die besseren Modellfit-Werte aufweisen kann, spiegelt am besten die Daten wider und wird daher angenommen. Die vergleichenden Modellfit-Werte sowie die Modelltests für die Modelle Aa, Ab und Ac werden in Tabelle 24 dargestellt. Da sich das Modell Ad aus den Ergebnissen der ersten drei Modelle ergeben hat, wird dieses erst nachfolgend beschrieben.

Tabelle 24: Modellfitwerte und -tests zur vergleichenden Bewertung der 1PL-Modelle zu Kompetenz in Biologie

Modellfitwerte und Modelltests	Modell Aa: Globalkompetenz Biologie	Modell Ab: zweidimensionales Standardmodell	Modell Ac: Zweidimensionales Zufallsmodell	Modell Ad:
<i>Final deviance</i>	70 513.00	70 488.36	70 554.56	70 493.02
<i>Loglikelihood</i>	-35 256.49	-35 244.18	-35 277.29	-35 246.51
AIC	71 056.99	71 034.36	71 100.57	71 037.02
BIC	72 705.28	72 688.71	72 754.93	72 685.32
Parameter	272	273	273	272
<i>Deviance change</i>		-24.64	41.56	-19.98
Satorra-Bentler χ^2 -Differenzentest		.00		
Wald-Test		19.55(1), .00	0.74(1), .39	

Anmerkungen. Da der Wald-Test nur für die Testung der vorgefundenen Korrelation der latenten Faktoren gegen eine Korrelation von Eins verwendet wurde und im Modell Ad keine Korrelation zwischen den latenten Faktoren angenommen wird, liegt für dieses Modell keine Prüfgröße des Wald-Tests vor. Die Werte des Wald-Tests werden wie folgt wiedergegeben: Prüfgröße(df), p-Wert. Der χ^2 -Differenzentest kann nur für genestete Modelle verwendet werden. Da das Modell Aa und Ac sowie Modell Aa und Ad nicht ineinander genestet sind, kann dieser für diese beiden Modelle nicht berechnet werden.

Wie aus Tabelle 24 ersichtlich wird, deuten alle Modellfit-Indizes und Modelltests darauf hin, dass das theoretisch hergeleitete zweidimensionale Modell Ab im Vergleich zu den beiden anderen Modellen am besten auf die Datenstruktur passt. Für das Modell Ab ist die *final deviance* am geringsten. Gegenüber dem eindimensionalen Modell verbessert sich die *final deviance* um 24.64, für das Zufallsmodell ist gegenüber dem eindimensionalen Modell eine Verschlechterung um 41.56 zu verzeichnen. Auch AIC und BIC sind für das Modell Ab am geringsten. Die Verbesserung zum eindimensionalen Modell beträgt für das $AIC_{\text{Change}} = 22.63$ und für das $BIC_{\text{Change}} = 16.57$ (Zufallsmodell: $AIC_{\text{Change}} = -43.58$ und $BIC_{\text{Change}} = -49.65$). Der Wald-Test bestätigt für das zweidimensionale Modell eine statistisch signifikant von Eins verschiedene Korrelation beider modellimplizierter Dimensionen. Im Zufallsmodell Ac ist die Korrelation beider Dimensionen nicht statistisch signifikant von Eins verschieden. Der Satorra-Bentler χ^2 -Differenzentest bestätigt dieses Bild. Auch diesem Test zufolge passt das Modell Ab statistisch signifikant besser auf die Daten als das eindimensionale Modell. Da das Zufallsmodell Ac nicht in dem eindimensionalen Modell genestet ist, ist ein analoger Vergleich mit dem Satorra-Bentler χ^2 -Differenzentest nicht möglich.

Aufgrund der Modellfit-Indizes wird das zweidimensionale Modell Ab favorisiert und die Forschungshypothese A bestätigt. Kompetenz in Biologie, wie sie mit dem vorliegenden Test

gemessen wird, ist somit zweidimensional strukturiert. Für das zweidimensionale Modell wird in der folgenden Abbildung 21 der Zusammenhang zwischen dem Konzept- und Prozesswissen abgebildet.

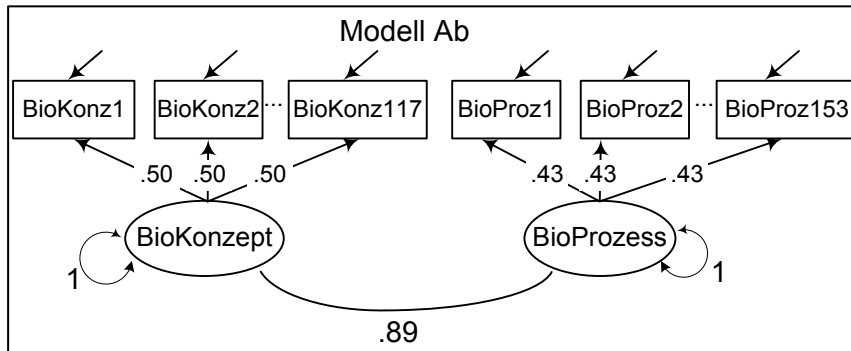


Abbildung 21: Zusammenhang der Dimensionen Konzept- und Prozesswissen von Kompetenz in Biologie

Anmerkungen. Bio = Biologie, Konz = Konzeptwissen, Proz = Prozesswissen

Die hohe Korrelation von $r = .89$ ($SE = .03$) zeigt, dass die beiden Dimensionen Konzept- und Prozesswissen in hohem Maße miteinander zusammenhängen. Die analoge Korrelation im Zufallsmodell Ac beträgt $r = .99$ ($SE = .02$). Dieser hohe Zusammenhang der zwei empirisch gefundenen Dimensionen wirft die Frage auf, ob die beiden Dimensionen nicht doch als ein gemeinsames Konstrukt abgebildet werden können. Daher wurde a posteriori das Modell Ad spezifiziert. In diesem Modell laden die manifesten Variablen zum Konzeptwissen auf einem latenten Faktor BioKonzept. Auf einem zweiten latenten Generalfaktor laden sowohl die manifesten Variablen zum Konzept- als auch zum Prozesswissen. Es handelt sich um ein *nested-factor* Modell, in welchem der Faktor Biologie Konzeptwissen in einem Generalfaktor genestet ist. In diesem Modell werden sowohl die generellen als auch die spezifischen Fähigkeiten der Schülerinnen und Schüler in einem Modell berücksichtigt (Gustafsson & Balke, 1993). Die beiden Faktoren korrelieren nicht miteinander. Dieses Modell weist auch einen guten Modellfit auf. Die Verbesserung des Modells Ad gegenüber dem eindimensionalen Modell Aa fällt jedoch nicht so hoch aus wie die Verbesserung von dem Modell Ab gegenüber dem Modell Aa. Auch das Modell Ad zeigte keine bessere Passung auf die Daten. Es sprechen somit zwei empirische Argumente für die Annahme der Zweidimensionalität. Zum einen passt das zweidimensionale Modell Ab besser auf die Daten als das eindimensionale Modell Aa und das Zufallsmodell Ac. Zum anderen fällt der Zusammenhang der zwei zufälligen Faktoren des Modells Ac mit $r = .99$ deutlich höher aus – liegt also nahe Eins – und die Modelltests favorisieren das zweidimensionale Modell. Auch wenn die zwei Dimensionen hoch miteinander korrelieren, muss also davon ausgegangen werden, dass die Zweidimensionalität ein besseres Abbild von Kompetenz in Biologie liefert.

5.1.2 Forschungshypothesen B1 sowie B2a und B2b: Zusammenhang von Kompetenz in Biologie mit verwandten kognitiven Konstrukten – Differentialpsychologische Perspektive I

Nach der Untersuchung der internen Struktur von Kompetenz in Biologie werden in einem zweiten Schritt den beiden Dimensionen Konzept- und Prozesswissen disziplinübergreifende Konstrukte gegenübergestellt. Die Zusammenhänge mit diesen Konstrukten können weitere Evidenzen für die Zweidimensionalität der Kompetenz in Biologie geben. In das auf Grundlage der Forschungshypothese A ermittelte Modell Ab werden in den folgenden Modellen die Konstrukte sprachliche Kompetenzen und kognitive Grundfähigkeit aufgenommen, um die Beziehung der beiden Dimensionen von Kompetenz in Biologie zu ihnen verorten zu können. Das Modell Ab wurde mit der ersten Forschungshypothese angenommen. Trotz des hohen Zusammenhangs zwischen den beiden Dimensionen ist es möglich, dass diese differentielle Zusammenhänge zu externen Kriterien aufweisen.

Zunächst sollen die Korrelationen der beiden Dimensionen von Kompetenz in Biologie mit den beiden disziplinübergreifenden Konstrukten in Tabelle 25 einen ersten Einblick in die differentiellen Zusammenhänge geben.

Tabelle 25: Messfehlerbereinigte Korrelationen zwischen den Dimensionen Konzept- und Prozesswissen mit den disziplinübergreifenden Konstrukten sprachliche Kompetenzen und kognitive Grundfähigkeit

Übergreifende Konstrukte	Biologie Konzeptwissen	Biologie Prozesswissen
Sprachliche Kompetenzen	.66	.65
Kognitive Grundfähigkeit	.57	.65

Es wird deutlich, dass ein differentielles Korrelationsmuster nur für die Dimension Konzeptwissen vorliegt. Die kognitive Grundfähigkeit hängt geringer mit dieser Teilkompetenz zusammen als die sprachlichen Kompetenzen. Im Prozesswissen spielen beide disziplinübergreifenden Konstrukte eine gleichbedeutende Rolle. Dieses Muster gibt erste Hinweise auf die möglichen differentiellen Zusammenhänge der beiden disziplinübergreifenden Konstrukte mit den beiden Kompetenzdimensionen.

Modell B: *Kompetenz in Biologie und kognitive Grundfähigkeit sowie sprachliche Kompetenzen*

Im folgenden Modell B wird das Konzept- und Prozesswissen auf die beiden latenten Faktoren sprachliche Kompetenzen und kognitive Grundfähigkeit regrediert. Die zwei latenten Faktoren

sprachliche Kompetenzen und kognitive Grundfähigkeit werden mit den Einzelvariablen der jeweiligen Tests in das Modell aufgenommen. Die kognitive Grundfähigkeit wird also von 30, die sprachlichen Kompetenzen von 28 manifesten Indikatoren gemessen. In Abbildung 22 wird das Modell mit den Ergebnissen abgebildet.

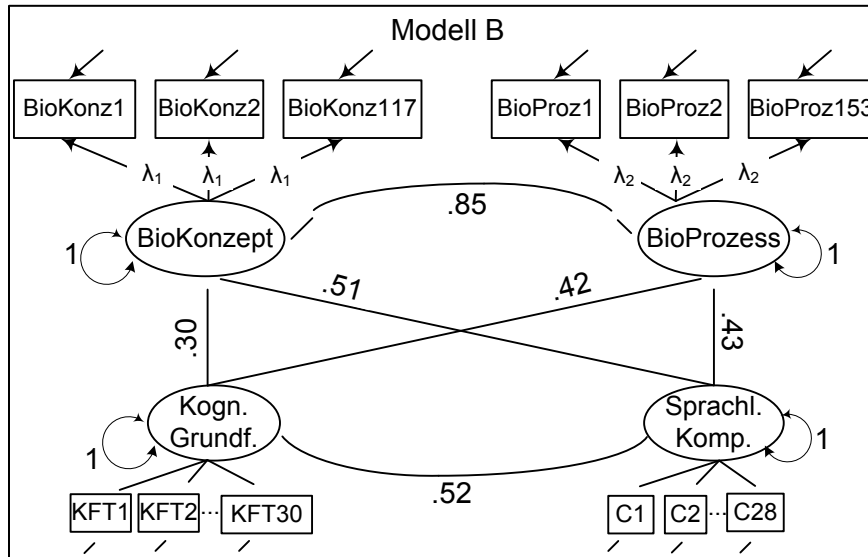


Abbildung 22: Ergebnisse des Strukturmodells zu den Forschungshypothesen B1, B2a und B2b – Konzept- und Prozesswissen und sprachliche Kompetenzen sowie kognitive Grundfähigkeit

Anmerkungen. Bio = Biologie, Konz = Konzeptwissen, Proz = Prozesswissen, Kogn. Grundf. = Kognitive Grundfähigkeit, Sprachl. Komp. = sprachliche Kompetenzen.

Die Korrelation zwischen sprachlichen Kompetenzen und kognitiver Grundfähigkeit liegt bei $r = .52$, die Residualkorrelation zwischen Konzept- und Prozesswissen bei $r = .85$. Diese Veränderung zur modellimplizierten Korrelation zwischen den beiden Faktoren liegt innerhalb des Intervalls von zwei Standardfehlern. Die modellimplizierten Regressionen und Korrelationen sind auf einem Signifikanzniveau von $\alpha = .01$ statistisch signifikant von Null verschieden. Die Forschungshypothese B1 kann auf Grundlage dieser Ergebnisse bestätigt werden. Es konnte gezeigt werden, dass die beiden disziplinübergreifenden Konstrukte einen Beitrag zur Erklärung der Dimensionen Konzeptwissen und Prozesswissen in Biologie leisten. Durch die sprachlichen Kompetenzen und die kognitive Grundfähigkeit werden 51% Varianz der Leistungen der Schülerinnen und Schüler im Konzeptwissen und 55% Varianz der Leistungen im Prozesswissen aufgeklärt. Sie werden also nicht vollständig erklärt.

Bezogen auf die Forschungshypothesen B2a und B2b liegt ein stärker differenziertes Bild vor. Nach bloßem Augenmaß könnte nur die Forschungshypothese B2a angenommen werden. Der Regressionskoeffizient von Konzeptwissen in Biologie auf die sprachlichen Kompetenzen ist

$\beta_{\text{direkt}} = .51$ ($SE = .02$, $\beta_{\text{total}} = .67$), auf die kognitive Grundfähigkeit $\beta_{\text{direkt}} = .30$ ($SE = .02$, $\beta_{\text{total}} = .57$). Der Unterschied zwischen den Regressionskoeffizienten beträgt also .21. Für das Prozesswissen liegen die beiden Regressionsgewichte bei $\beta_{\text{LK(direkt)}} = .43$ ($SE = .02$, $\beta_{\text{LK(total)}} = .65$) und $\beta_{\text{kGf(direkt)}} = .42$ ($SE = .02$, $\beta_{\text{kGf(total)}} = .64$). Sprachliche Kompetenzen spielen in beiden Teilkompetenzen von Biologie eine größere Rolle als die kognitive Grundfähigkeit.

Zur Prüfung der beiden Forschungshypothesen wurden mit dem Wald-Test jeweils die Regressionskoeffizienten Prozesswissen auf sprachliche Kompetenzen und kognitive Grundfähigkeit sowie die Regressionskoeffizienten Konzeptwissen auf sprachliche Kompetenzen und kognitive Grundfähigkeit auf Gleichheit geprüft. Für das Konzeptwissen ist die Differenz beider Regressionskoeffizienten statistisch signifikant von Null verschieden; die Koeffizienten sind also nicht gleich ($\chi^2 = 19.86$; $df = 1$; $p < .00$). Somit hängen die sprachlichen Kompetenzen höher mit dem Konzeptwissen in Biologie zusammen als die kognitive Grundfähigkeit. Dies bedeutet, dass die Forschungshypothese B2a bestätigt wird. Für das Prozesswissen konnte jedoch kein statistisch signifikanter Unterschied dieser beiden Regressionen von einer Gleichsetzung gezeigt werden ($\chi^2 = 0.10$; $df = 1$; $p < .76$). Die kognitive Grundfähigkeit hängt also nicht höher mit dem Prozesswissen in Biologie zusammen als die sprachlichen Kompetenzen. Die Forschungshypothese B2b muss folglich abgelehnt werden. Hier sind die beiden Regressionsgewichte nicht statistisch signifikant voneinander verschieden. Dieses differentielle Regressionsmuster steht im Einklang mit den bereits berichteten korrelativen Zusammenhängen der disziplinübergreifenden Konstrukte mit den beiden Dimensionen.

Abschließend bleibt zu bemerken, dass ein gesonderter Blick auf die disziplinübergreifenden Konstrukte erwartbare Unterschiede verdeutlicht. So steht die kognitive Grundfähigkeit in einem geringeren Zusammenhang mit dem Konzeptwissen in Biologie als mit dem Prozesswissen. Der Unterschied beträgt .12. Die sprachlichen Fähigkeiten hängen mit dem Konzeptwissen in Biologie wiederum höher zusammen als mit dem Prozesswissen. Hier beträgt der Unterschied der beiden Regressionskoeffizienten .08. Es konnten somit differentielle Effekte der einzelnen disziplinübergreifenden Konstrukte auf das Konzept- und Prozesswissen in Biologie gezeigt werden.

5.1.3 Forschungshypothesen C und D: Zusammenhang von Kompetenz in Biologie mit Fachnoten und naturwissenschaftlichen Selbstkonzepten – Differentialpsychologische Perspektive II

Mit den beiden folgenden Forschungshypothesen werden nun weitere Kriterien in die Strukturmodelle aufgenommen, von denen angenommen wird, dass im Vergleich einige dem Konstrukt Kompetenz in Biologie näher stehen und andere geringer mit dem Konstrukt im Zusammenhang stehen. Zum einen sollten Fachnoten unterschiedlicher Disziplinen unterschiedliche Zusammenhänge zum Kompetenztest zeigen. Zum anderen sollten Selbstkonzepte anderer

naturwissenschaftlicher Disziplinen dem Konstrukt Kompetenz in Biologie weniger nahe stehen als das Selbstkonzept in Biologie. Zunächst sollen die Fachnoten in das nomologische Netz der Validierung aufgenommen werden. Die Korrelationen der Fachnoten mit den beiden Dimensionen von Kompetenz in Biologie sind in Tabelle 26 abgebildet.

Tabelle 26: Messfehlerbereinigte Korrelationen zwischen den Dimensionen Konzept- und Prozesswissen und den Fachnoten in Deutsch und Mathematik sowie den drei naturwissenschaftlichen Fächern Biologie, Chemie und Physik

Fachnoten	Biologie Konzeptwissen	Biologie Prozesswissen
Deutsch	.26	.26
Mathematik	.24	.24
Biologie	.26	.25
Chemie	.22	.23
Physik	.25	.26

Auf den ersten Blick fällt besonders auf, dass die Fachnoten in allen Bereichen ähnlich hoch mit den beiden Dimensionen von Kompetenz in Biologie korrelieren, Unterschiede zwischen den Fachnoten sind kaum auszumachen. Mit dem Konzeptwissen hängen die Fachnoten in Deutsch und Biologie am höchsten zusammen, mit dem Prozesswissen die Fachnoten in Deutsch und Physik. Die Unterschiede sind allerdings zu gering, um Hinweise für die modellbasierten Rechnungen zu geben.

Modell C: Zusammenhänge von Kompetenz in Biologie mit Fachnoten in Biologie, Chemie, Physik, sowie Deutsch und Mathematik

In dem Modell C werden die Zusammenhänge der Dimensionen Konzept- und Prozesswissen in Biologie mit den Noten in den Schulfächern Biologie, Physik, Chemie, Deutsch und Mathematik modelliert. Die am Klassenmittelwert zentrierten Fachnoten gehen als manifeste Prädiktoren in das Modell ein und beschreiben die beiden latenten Faktoren Biologie Konzept- und Prozesswissen. Da die klassenzentrierten Fachnoten in das Modell eingingen, wurde für dieses Modell die Schule als Clustervariable aufgenommen. Das Modell ist in Abbildung 23 dargestellt.

Ein methodisches Problem tritt bei der vergleichenden Analyse von Noten unterschiedlicher Fächer auf. Im Gegensatz zu den bisher verwendeten Maßen wurde der Schülerin bzw. dem Schüler jede Note von einer anderen Lehrkraft gegeben. Eine Schülerin X hat z. B. ihre Fachnoten in Biologie und Chemie von der gleichen Lehrkraft, in den Schulfächern Physik, Mathematik und Deutsch von unterschiedlichen Lehrkräften erhalten. Diese unterschiedlichen Lehrkräfte legen ihrer Benotung

auch unterschiedliche Benotungssysteme zu Grunde. In Hinblick auf diese Unterschiede zwischen den Fachnoten sollten auch die Unterschiede in den Noten mit Vorsicht interpretiert werden.

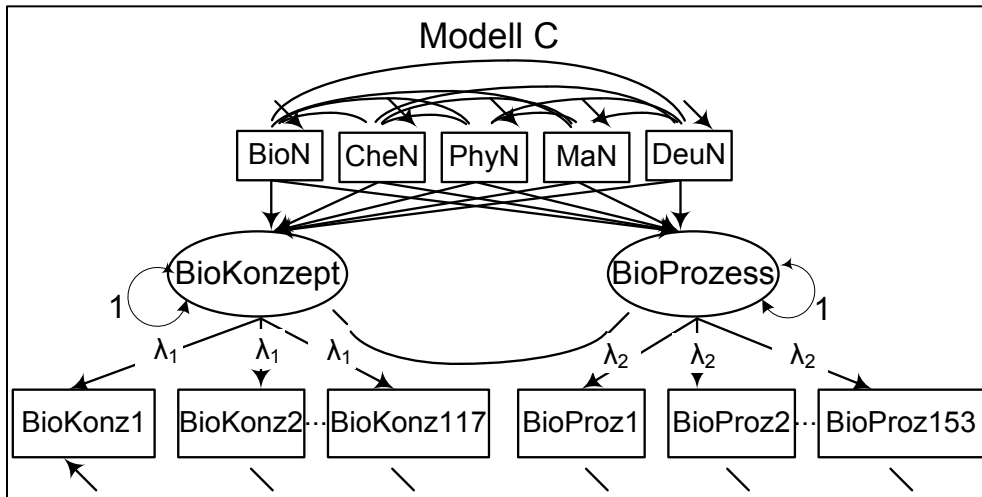


Abbildung 23: Spezifikation des Strukturmodells zu den Forschungshypothesen C1, C2 und C3 – Kompetenzen in Biologie und Fachnoten in Biologie, Chemie und Physik sowie in Mathematik und Deutsch

Anmerkungen. Bio = Biologie, Che = Chemie, Phy = Physik, Ma = Mathematik, Deu = Deutsch, N = Noten, Konz = Konzeptwissen, Proz = Prozesswissen.

Die naturwissenschaftlichen Fachnoten korrelieren untereinander zwischen $r = .52$ und $r = .59$. Die naturwissenschaftlichen Fachnoten korrelieren mit den Fachnoten in Deutsch und Mathematik zwischen $r = .42$ und $r = .60$. Die Fachnoten Deutsch und Mathematik korrelieren untereinander zu $r = .39$. Die einzelnen Korrelationen können der Tabelle 36 im Anhang entnommen werden. Die Residualkorrelation zwischen den beiden Dimensionen von Kompetenz in Biologie beträgt $r = .88$. In Tabelle 27 werden die Regressionskoeffizienten der Fachnoten auf die beiden Teildimensionen dieses Modells aufgelistet.

Tabelle 27: Regressionsgewichte der Fachnoten Biologie, Chemie, Physik, Mathematik und Deutsch für die beiden Teildimensionen Konzept- und Prozesswissen in Biologie und deren Standardfehler

Schulfach	Konzeptwissen		Prozesswissen	
	β	SE	β	SE
Deutsch	.13*	.03	.13*	.03
Mathematik	.08	.04	.09	.03
Biologie	.10	.03	.08	.03
Chemie	.01	.03	.02	.04
Physik	.09	.04	.10	.04

Anmerkungen. β = Standardisiertes Regressionsgewicht, SE = Standardfehler

* = Regressionsgewichte sind auf einem Niveau von $\alpha = .05$ statistisch signifikant von Null verschieden.

Für beide Teilbereiche weist die Fachnote in Deutsch die höchste Regression auf. Für das Konzeptwissen in Biologie folgt die Fachnote in Biologie. Im Prozesswissen ist die nächstniedrige Regression auf die Fachnote Physik, erst dann folgt die Fachnote in Biologie. Prinzipiell sind die Regressionsgewichte sehr niedrig und liegen nah beieinander. Die Fachnote in Biologie weist also für beide Kompetenzbereiche nicht das höchste Regressionsgewicht auf. Im Konzeptwissen ist die Fachnote in Biologie – wenn auch mit geringem Abstand – die höchste innerhalb der naturwissenschaftlichen Disziplinen, im Prozesswissen wird dieses Bild weitestgehend repliziert. Somit müssen die Forschungshypothesen C1, C2 und C3 abgelehnt werden. Die erhobenen Teildimensionen von Kompetenz in Biologie stehen nicht in einem mittleren Zusammenhang mit den Fachnoten in Biologie und der Zusammenhang mit den Fachnoten in den weiteren naturwissenschaftlichen Disziplinen fällt nicht bedeutend geringer als der Zusammenhang der Fachnote in Biologie aus. Vielmehr wird über die naturwissenschaftlichen Fachnoten hinweg ein sehr einheitliches Muster der Regressionsgewichte deutlich. Die Regression der beiden nicht-naturwissenschaftlichen Disziplinen fügt sich in dieses Muster ein, sodass auch hier die Regressionskoeffizienten nicht geringer ausfallen als die der naturwissenschaftlichen Fächer. Vielmehr ist es das Gegenteil der Fall, in beiden Teildimensionen überwiegt die Note im Fach Deutsch. Da dieses Regressionsgewicht nur geringfügig höher als in den anderen Fächern ausfällt, sollte dieser Befund nicht überbewertet werden. Insgesamt klären die Fachnoten der fünf Disziplinen jeweils 11% der Varianz der latenten Faktoren Konzept- und Prozesswissen auf. Um zu zeigen, wie viel Varianz in diesem Modell die Fachnote in Biologie in den beiden latenten Faktoren Konzept- und Prozesswissen aufklärt, wurde ein Modell spezifiziert, in dem die Fachnote Biologie nicht mit aufgenommen wurde. Dieses Modell zeigt, dass von der Fachnote in Biologie jeweils nur weniger als

1 % im Konzept- und Prozesswissen aufgeklärt werden. Das Modell ohne die Fachnote in Biologie kann der Tabelle 37 im Anhang entnommen werden.

Modell D: Zusammenhänge von Kompetenz in Biologie mit naturwissenschaftlichen Selbstkonzepten

Zur Klärung der Forschungshypothese D wurden zwei weitere Strukturmodelle zu dem Persönlichkeitsmerkmal Selbstkonzept spezifiziert. Aufgrund der hohen Anzahl an manifesten Variablen und fehlenden Werten konnte diese Hypothese nur getrennt für die beiden Teilkompetenzen Konzept- (Modell Da) und Prozesswissen (Modell Db) getestet werden. Das Gesamtmodell konvergierte nicht. Die korrelativen Zusammenhänge (s. Tabelle 28) zeigen, dass das Selbstkonzept in Biologie mit beiden Dimensionen höher zusammenhängt als das Selbstkonzept in Chemie und Physik.

Tabelle 28: Messfehlerbereinigte Korrelationen zwischen den Dimensionen Konzept- und Prozesswissen mit den Selbstkonzepten in Biologie, Chemie und Physik

Selbstkonzept	Biologie Konzeptwissen	Biologie Prozesswissen
Biologie	.26	.25
Chemie	.14	.12
Physik	.16	.12

Die Tabelle macht deutlich, dass das Selbstkonzept in Biologie mit beiden Dimensionen von Kompetenz in Biologie höher korreliert als das Selbstkonzept in Chemie oder in Physik. In den beiden Modellen Da und Db (s. Abbildung 24) zu den naturwissenschaftlichen Selbstkonzepten werden die Selbstkonzepte der Schülerinnen und Schüler in den Disziplinen Biologie, Chemie und Physik als latente Faktoren spezifiziert. Als manifeste Variablen für die drei latenten Selbstkonzeptfaktoren werden wiederum die Einzelitems der Skala herangezogen. Die Dimensionen Konzept- und Prozesswissen werden jeweils auf die latenten Selbstkonzeptfaktoren regrediert.

Modell Da: Konzeptwissen in Biologie und Selbstkonzepte in den naturwissenschaftlichen Disziplinen. In diesem Modell wird das Konzeptwissen auf die Selbstkonzepte in Biologie, Chemie und Physik regrdiert.

Modell Db: Prozesswissen in Biologie und Selbstkonzepte in den naturwissenschaftlichen Disziplinen. In diesem Modell wird das Prozesswissen auf die Selbstkonzepte in Biologie, Chemie und Physik regrediert.

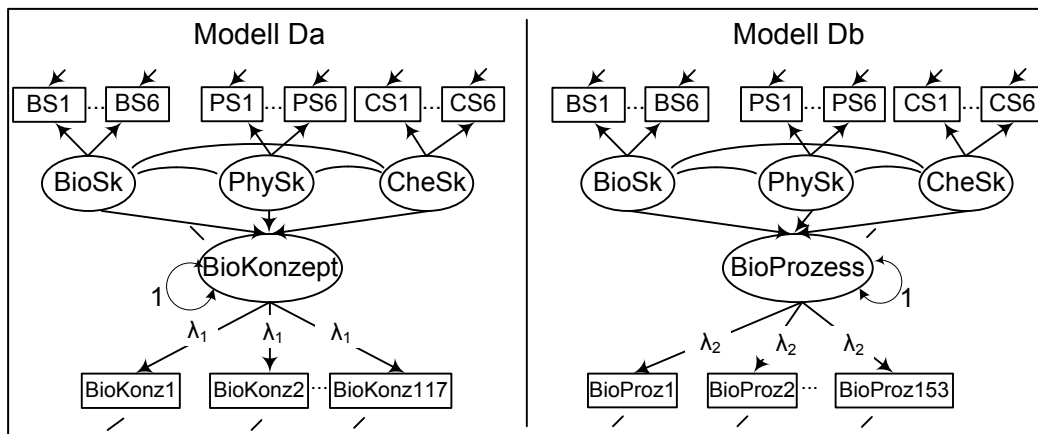


Abbildung 24: Spezifikation der Strukturmodelle zur Forschungshypothese D – Naturwissenschaftliche Selbstkonzepte und Konzept- und Prozesswissen in Biologie

Anmerkungen. BS, BioSk = Biologie Selbstkonzept, PS1, PhySk = Physik Selbstkonzept, CS, CheSk = Chemie Selbstkonzept, Bio = Biologie, Konz = Konzeptwissen, Proz = Prozesswissen.

Zunächst wird das Modell Da näher beschrieben, also die Zusammenhänge der Selbstkonzepte mit der Teildimension Konzeptwissen. Die Selbstkonzepte untereinander korrelieren zu $r_{\text{Bio-Che}} = .47$, $r_{\text{Bio-Phy}} = .37$ und $r_{\text{Phy-Che}} = .56$. Der Regressionskoeffizient von Konzeptwissen in Biologie auf das akademische Selbstkonzept in Biologie weist mit $\beta_{\text{Biologie}} = .23$ ($SE = .03$) den höchsten Wert auf und ist statistisch signifikant von Null verschieden. Die weiteren beiden Regressionsgewichte liegen bei $\beta_{\text{Chemie}} = -.02$ und $\beta_{\text{Physik}} = .04$. Diese beiden Regressionsgewichte sind auch nicht statistisch signifikant von Null verschieden. Die drei naturwissenschaftlichen Selbstkonzeptfaktoren klären 6 % der Varianz des latenten Faktors Konzeptwissen auf.

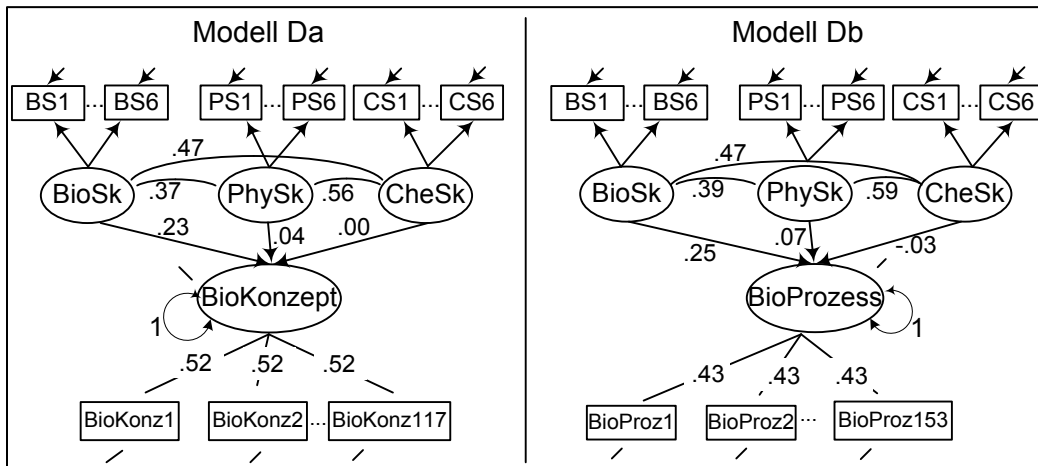


Abbildung 25: Ergebnisse der Strukturmodelle zu Forschungshypothese D – Naturwissenschaftliche Selbstkonzepte und Konzept- und Prozesswissen in Biologie

Anmerkungen. BS, BioSk = Biologie Selbstkonzept; PS, PhySk = Biologie Selbstkonzept; CS, CheSk = Chemie Selbstkonzept; Bio = Biologie; Konz = Konzeptwissen; Proz = Prozesswissen.

Das Modell Db repliziert dieses Bild für die Dimension Prozesswissen in Biologie. Die Selbstkonzepte korrelieren untereinander zu $r_{Bio-Che} = .47$, $r_{Bio-Phy} = .39$ und $r_{Phy-Che} = .59$. Das Selbstkonzept in Biologie trägt auch für diese Dimension am meisten zur Varianzaufklärung bei ($\beta_{Biologie} = .25$, $SE = .03$, $R^2_{Biologie} = .06$). Die weiteren beiden naturwissenschaftlichen Selbstkonzepte zeigen nicht statistisch signifikante Regressionsgewichte von $\beta_{Chemie} = -.03$ und $\beta_{Physik} = .08$. Durch die drei naturwissenschaftlichen Selbstkonzepte werden 7% der Varianz aufgeklärt. Damit tragen in beiden Modellen die akademischen Selbstkonzepte in Chemie und Physik nur minimal zur weiteren Varianzaufklärung bei.

Um die Varianzaufklärung des Selbstkonzepts in Biologie zu verdeutlichen, wurden wiederum Modelle spezifiziert, in denen das Selbstkonzept in Biologie nicht mit aufgenommen wurde. Diese Modelle zeigen, dass von dem Selbstkonzept in Biologie 3 % Varianz im Konzeptwissen aufgeklärt werden, dies entspricht der Hälfte der Gesamtvarianz. Im Prozesswissen werden 5 % der Varianz durch das Selbstkonzept in Biologie aufgeklärt. Dies entspricht sogar mehr als der Hälfte der gesamten Varianzaufklärung. Die beiden Modelle ohne das Selbstkonzept in Biologie können der Tabelle 38 im Anhang entnommen werden.

6 Diskussion

Im Zuge der Entwicklung von Bildungsstandards für verschiedene Kompetenzbereiche sowie der Ausrichtung des Bildungssystems auf Outputorientierung und Kompetenzansätze bekommt die Validität von internationalen und nationalen Kompetenztests immer mehr Bedeutung. Aufgrund ihrer Testergebnisse werden bildungspolitische Entscheidungen getroffen, die für zukünftige Schülergenerationen weitreichende Konsequenzen haben können. Mit der vorliegenden Arbeit wurden Aspekte der Validität von Kompetenz in Biologie aus der fachdidaktisch-kognitionspsychologischen und aus der differentialpsychologischen Perspektive bearbeitet. Ein besonderes Augenmerk wurde hierbei auf die inhaltliche Relevanz und Repräsentativität der Teilkompetenzen Konzept- und Prozesswissen und deren Zusammenhänge mit externen Variablen gesetzt. Beide Perspektiven führten zu einem differentiellen Ergebnismuster. Deskriptive Analysen zeigten, dass es gelungen ist, Testaufgaben unterschiedlicher Schwierigkeit zu entwickeln.

Die fachdidaktisch-kognitionspsychologische Perspektive – mit der die interne Struktur von Kompetenz in Biologie näher untersucht wurde – zeigte, dass die Testaufgaben Kompetenz in den beiden Teilkompetenzen Konzept- und Prozesswissen in Biologie zu messen scheinen. Die Teilkompetenzen wurden als hoch miteinander korrelierende Dimensionen herausgearbeitet. Diese Ausdifferenzierung fügt sich in den Rahmen der *scientific literacy* Diskussion und der sich daraus ergebenden theoretischen Modelle zur Dimensionalität von naturwissenschaftlicher Kompetenz. Auch die kognitionspsychologischen Theorien stehen mit diesem Ergebnis im Einklang.

Die differentialpsychologische Perspektive und der damit einhergehende Einbezug externer Kriterien gaben einen weiteren Einblick in diese zweidimensionale Struktur. Die Hinzunahme der Konstrukte sprachliche Kompetenzen und kognitive Grundfähigkeit zeigte, dass das Konzeptwissen mehr durch die sprachlichen Kompetenzen als durch die kognitive Grundfähigkeit determiniert wird. Das Prozesswissen wird hingegen zu gleichen Teilen von diesen beiden disziplinübergreifenden Fähigkeiten determiniert. Der Blick auf die Fachnoten in den drei naturwissenschaftlichen Fächern sowie in Deutsch und Mathematik konnte das zuvor gezeigte Bild vervollständigen. Die Deutschnote leistete den größten Beitrag in beiden Teildimensionen; im Konzeptwissen gefolgt von der Fachnote in Biologie, im Prozesswissen von der Fachnote in Physik. Die schulischen Fachleistungen klären insgesamt nur zu einem geringen Anteil die Varianz in den beiden Dimensionen auf. Auch das Selbstkonzept trug zu einem relativ geringen Teil zur Erklärung der beiden Teilkompetenzen bei. Es konnten jedoch die prognostizierten Regressionsmuster gezeigt werden. Die Resultate zur letzten Forschungshypothese zeigten, dass zur Förderung von *scientific literacy* und *biological literacy*

Persönlichkeitsaspekte wie das Selbstkonzept eine wichtige Rolle spielen. Das Selbstkonzept in Biologie steht mit beiden Teildimensionen von Kompetenz in Biologie in einem höheren Zusammenhang als die Selbstkonzept in Chemie und Physik. Diese verschiedenen Ergebnisse zur internen Struktur sollten bei der Unterrichtsgestaltung, bei der Durchführung von *large-scale assessment* Studien und für weitere wissenschaftliche Studien in Betracht gezogen werden. Wie dies realisiert werden kann, wird im Folgenden bezogen auf die vier Fragestellungen der vorliegenden Arbeit näher erläutert.

6.1 Forschungshypothese A: Zusammenhang des Konzept- und Prozesswissens in Biologie

Bezogen auf die interne Struktur von Kompetenz in Biologie wurden drei konkurrierende Modelle berechnet, ein eindimensionales Modell Aa, ein zweidimensionales Modell Ab und ein zufällig zweidimensionales Modell Ac. Das Modell Ab zeigte im Vergleich den besten Modellfit. Zudem zeigten die Modelltests Evidenz für dieses theoriegeleitete zweidimensionale Modell. Somit wird die Forschungshypothese A bestätigt; Kompetenz in Biologie besteht aus den zwei Subkompetenzen Konzept- und Prozesswissen. Erste Hinweise für diese Teilbereiche von Kompetenz in Biologie konnten in den Naturwissenschaften (Senkbeil et al., 2005) und in der Chemie (Klos et al., 2008) gezeigt werden und können mit der vorliegenden Untersuchung für die Biologie bestätigt werden. Dieses Ergebnis steht des Weiteren im Einklang mit kognitionspsychologischen Auffassungen von Wissensstrukturen (z. B. Anderson, 1983; Edelman, 2000; Lukesch, 2001). Die beiden vorgefundenen Subkompetenzen korrelieren jedoch sehr hoch miteinander. Dieser Befund fügt sich in die Ergebnisse aus den PISA-Studien ein, liegt jedoch höher als in Zusatzanalysen der FIMS- und SISS-Studie (Tamir et al., 1992). Der Zusammenhang fällt höher aus als in einigen vorangegangenen Studien (Klos et al., 2008; OECD, 2009). Er befindet sich ungefähr auf dem Niveau der Korrelation der naturwissenschaftlichen Fächer untereinander (Senkbeil et al., 2005). Zunächst wird die Frage diskutiert, welche Gründe und Implikationen eine zweidimensionale Kompetenz in Biologie haben könnte.

Damit Schülerinnen und Schüler Strategien einer akademischen Disziplin anwenden können, müssen sie über eine gewisse Wissensbasis in verschiedenen Teilbereichen dieser Disziplin verfügen. Die Bildungsstandards für das Fach Biologie explizieren daher Kompetenzen in den Bereichen Konzept- und Prozesswissen auf nationaler Ebene als wichtige Ziele von Unterricht (KMK, 2005a). Alle Schülerinnen und Schüler sollten demnach ein Verständnis über die Konzepte und Methodik der Naturwissenschaften erlangen (z. B. Hodson, 2003; Shen, 1975). Diese könnten auch im späteren Berufsleben von den Schülerinnen und Schülern gefordert werden (Gott & Duggan, 1999).

Vorangegangene Studien zeigten jedoch, dass das Prozesswissen sowohl unter naturwissenschaftlich ausgerichteten Forscherinnen und Forschern (Wong & Hodson, 2009) als auch im Zusammenhang mit Unterrichten in den naturwissenschaftlichen Fächern (Roberts & Gott, 2000) eher als *common sense* denn als eigenständiger Wissensbestand angesehen wird, welcher explizit gelehrt werden müsste. Die vorliegenden Ergebnisse machen jedoch deutlich, dass ein eigenständiger Wissensstand vorzuliegen scheint, der eng mit dem Konzeptwissen verzahnt ist. Das vorrangige Unterrichten von Konzeptwissen und das stark angeleitete Experimentieren werden der hier vorgefundenen Mehrdimensionalität von Kompetenz in den Naturwissenschaften nicht gerecht. Es kann weder den geforderten Anwendungsbezug (Bybee, 2002) leisten, noch fokussiert es auf das Konzept und die Ziele der *scientific literacy* (DeBoer, 2000). Unterricht, der Prozesswissen in seiner Breite und Tiefe nicht ausreichend berücksichtigt, ist der kulturellen Teilhabe der Schülerinnen und Schüler in ihrem jetzigen und zukünftigen Leben nicht dienlich, da ihnen der Modus der kognitiv-instrumentellen Modellierung der Welt als einer der Weltbegegnungen (Baumert, 2002) teilweise verwehrt bleibt.

Der hohe Zusammenhang zwischen den beiden Teilkompetenzen Konzept- und Prozesswissen deutet darauf hin, dass Schülerinnen und Schüler beim konzeptuellen Lernen auch Wissen zu den Prozessaspekten der Biologie bzw. in den Naturwissenschaften erlangen und *vice versa* (Hodson, 1992). Lehrkräfte berücksichtigen folglich bei der Vermittlung von Prozessen implizit auch die dahinter verborgenen Konzepte und bei der Vermittlung von Konzepten die damit einhergehenden Prozesse. Zudem ist es möglich, dass einige Schülerinnen und Schüler eher von der Vermittlung von Konzeptwissen profitieren, andere Schülerinnen und Schüler eher von der Vermittlung von Prozesswissen. So könnte das Wissen in Biologie die Wahl eines experimentellen Designs beeinflussen (Zimmerman, 2007). Die differentiellen Zusammenhänge während des Wissenserwerbs können nur anhand von Längsschnitt- oder Interventionsstudien quasiexperimentellen Charakters untersucht werden, wie sie für die Mathematik bereits vorliegen (s. Seite 58). Das vorliegende Ergebnis kann für die Erstellung von Kompetenzentwicklungsmodellen eine Grundlage bilden. Für die Mathematik konnte eine wechselseitige Beeinflussung innerhalb des Wissenserwerbsprozesses aufgezeigt werden. Aus diesen Interventions- und Längsschnittuntersuchungen können Kompetenzentwicklungsmodelle entstehen, welche die gegenseitige Beeinflussung beim Kompetenzerwerb aufschlüsseln. Um diese entstandenen Modelle für den Unterricht nutzbar zu machen, sollten auf deren Grundlage Lernaufgaben entwickelt werden (Bernholt et al., 2009).

Unterricht wird durch bundesländerspezifische Curricula gesteuert. In einigen Bundesländern wurden die Curricula mit dem Ziel der Implementation der Bildungsstandards bereits überarbeitet (z. B. Berlin), in anderen steht diese Überarbeitung noch aus (z. B. Bayern). Auch an dieser Stelle sollten alle Teilaspekte von Kompetenz in Biologie berücksichtigt werden (Roberts & Gott, 2004). Um

ein Curriculum zu entwerfen, das die Biologie mit all ihren Aspekten widerspiegelt, muss die Biologie zunächst mit all ihren Aspekten definiert werden. In einem zweiten Schritt können dann die Unterrichtsinhalte in allen Wissensarten selektiert werden. Nur so wird die Biologie der „realen Welt“ in die Schulen gebracht und kann im späteren Leben angewandt werden. Prozesswissen sollte also als eine eigene Komponente und in seiner Differenziertheit in die Curricula aufgenommen werden. Bei der momentanen Umstellung der Curricula, könnte der gefundene hohe Zusammenhang zwischen den beiden Kompetenzbereichen berücksichtigt werden. Für eine tatsächliche Umsetzung von intendierten Curricula werden zusätzlich Änderungen im Lehrerwissen und der Lehrerprofessionalität benötigt (Hodson, 1992). Gleichzeitig sollte die enge Verzahnung, die der hohe Zusammenhang vermuten lässt, im Unterricht berücksichtigt werden. Implizites Lernen durch die Praxis gelingt nur den leistungsstarken Schülerinnen und Schülern (Gott et al., 1999). Das explizite Lehren von Erkenntnisgewinnung verdoppelt den Gewinn (Roberts, 2001; Tamir et al., 1998). Neben der Verzahnung ist eine weitere Herausforderung, das Gleichgewicht zwischen naturwissenschaftlichen Inhalten und den weiteren wichtigen Zielen naturwissenschaftlicher Bildung zu finden (DeBoer, 2000). Auch die Struktur von nationalen und internationalen Leistungserhebungen muss einem umfassenden Curriculum gerecht werden und die Biologie in ihrer Ganzheit abbilden. Mit der Erhebung von Prozesswissen in Biologie als einer eigenständigen Testdimension in schriftlicher Form ist ein erster Schritt in diese Richtung getan (Roberts & Gott, 2004).

Die schriftliche Testung von Prozesswissen, wie sie in der vorliegenden Untersuchung angewendet wurde, birgt das Problem in sich, dass nur ein Teil des Konstrukts Prozesswissen getestet werden kann (Shavelson et al., 1991). Das tatsächliche Handeln könnte nur über praktisches Testen approximiert werden. Diese praktischen Tests bringen jedoch Reliabilitätsprobleme mit sich. Daher ist für beide Testformen bei der Übertragung der Ergebnisse in die Realität Vorsicht geboten. Dennoch birgt eine breit angelegte Studie zu naturwissenschaftlicher Kompetenz die Möglichkeit, Unterschiede von Subpopulationen, z. B. von Mädchen und Jungen oder von Schülerinnen und Schülern unterschiedlicher Schulformen, für unterschiedliche Teilkompetenzen zu untersuchen.

Mit dem Ziel einer umfassenden Bildung von naturwissenschaftlicher Kompetenz (Strobert, 2009) sollte auch eine entsprechende umfassende Testung der naturwissenschaftlichen Leistungen von Schülerinnen und Schülern einhergehen (Harlow & Jones, 2004). In Folgeuntersuchungen sollte in Erwägung gezogen werden, zusätzlich praktische Fähigkeiten, also z. B. das aktive Experimentieren oder das Beobachten im Feld, für Teilstichproben zu erheben. Die Hinzunahme der alternativen Erhebungsmethoden würde es ermöglichen, auch den Prozess abzubilden, der zu den Wissensbeständen der Schülerinnen und Schülern führt (Baartman et al., 2007). Validitäts- und

Reliabilitätsprobleme sollten von einer flächendeckenden Testung geprüft und so weit wie möglich vermieden werden. Zu dem hier angewendeten schriftlichen Test wäre z. B. eine praktische Testung mittels Experimentier- und Beobachtungsaufgaben möglich.

Die vorliegende Studie hat die Zweidimensionalität von Kompetenz in Biologie offen gelegt. Hieran sollten weitere Untersuchungen angeschlossen werden, die die interne Struktur innerhalb des Prozesswissens untersuchen. So könnten die differentiellen Fähigkeiten der Schülerinnen und Schüler durch spezifische Förderstrategien verbessert werden. So wie das Konzeptwissen aus unterschiedlichen Konzepten zusammengesetzt ist, wird auch das Prozesswissen differenziert werden können. Hierzu liegen erste Ergebnisse vor, die Hinweise darauf geben, in welchen Teilkompetenzen Schülerinnen und Schüler Förderung benötigen (Mayer et al., 2008). Informationen zu diesen Teilkompetenzen könnten Lehrerinnen und Lehrer bei der Schwerpunktsetzung in ihrer Unterrichtsgestaltung unterstützen, um einer einseitigen Vermittlung von Teilaspekten der naturwissenschaftlichen Erkenntnisgewinnung (z. B.: Kontrolle von Variablen in Laborexperimenten) vorzubeugen. Bei der Vermittlung sind zwei Schwerpunktsetzungen von Bedeutung. Zum einen sollten Lehrerinnen und Lehrer das gesamte Repertoire an Herangehensweisen berücksichtigen. Zum anderen sollte auch der gesamte Prozess einer jeden Herangehensweise der Erkenntnisgewinnung von der Auswahl der adäquaten Methode bis zur Auswertung und Interpretation der Daten von den Schülerinnen und Schülern nachempfunden werden. Nur so werden sie dazu befähigt, in einer Situation ihres alltäglichen Lebens naturwissenschaftliche Methoden anzuwenden. Schließlich wurden die beiden Subkompetenzen Konzept- und Prozesswissen untersucht, also nur zwei – wenn auch bedeutende – Teilbereiche von Kompetenz in Biologie unter vielen möglichen. So explizieren die Bildungsstandards die beiden weiteren Kompetenzbereiche Bewertung und Kommunikation. Untersuchungen zu weiteren Teilkompetenzen müssen in der Zukunft angestrebt werden, um ein umfassenderes Bild von der Struktur von Kompetenz in Biologie zu erhalten.

Die hohe Korrelation der beiden Teildimensionen von Kompetenz in Biologie ist u. a. auf das Zusammenspiel der beiden Wissensarten Konzept- und Prozesswissen in aktuellen Situationen zurückzuführen (Anderson, 2001). Die Nutzung und Erlangung von Wissen in einer Dimension scheint somit positiv auf das Wissen in der anderen Dimension zu wirken. Des Weiteren zeigt der Forschungsstand der Kognitionspsychologie, dass Strategien durch die Prozeduralisierung des deklarativen Wissens entstehen (Anderson, 2001). Es ist also möglich, dass auch disziplinspezifische Strategien wie das Durchführen von Experimenten aus der Prozeduralisierung von disziplinspezifischem deklarativem Wissen entstehen. Der hohe Zusammenhang von $r = .89$ könnte dann damit erklärt werden, dass die eine Wissensart Prozesswissen eine neue, erweiterte Form der

anderen Wissensart Konzeptwissen darstellt. Zudem werden sowohl das Konzept- als auch das Prozesswissen anhand von Inhalten – also Konzepten – abgerufen.

Der hohe Zusammenhang zwischen den beiden Dimensionen von Kompetenz in Biologie könnte zudem zu einem nicht bestimmbareren Anteil auf einen Methodeneffekt zurückzuführen sein. Sowohl das Konzept- als auch das Prozesswissen wurden über einen schriftlichen Test abgefragt. Frühere Untersuchungen haben gezeigt, dass die Korrelationen zwischen zwei schriftlichen Tests höher sind, als zwischen einem schriftlichen und einem praktischen Test (Tamir & Doran, 1992). Somit könnte die „wahre“ Korrelation zwischen Konzept- und Prozesswissen geringer ausfallen als die hier berichtete. Des Weiteren kann ein schriftlicher Test nicht die Vorgänge einer *hands-on* Aktivität und daher nicht das Handeln einer Testperson in einer alltagsnahen naturwissenschaftlichen Situation ersetzen. Praktische Tests wurden in einer deutschlandweiten Studie wie der vorliegenden bisher nicht durchgeführt. Daher liegt auch nur eingeschränktes Wissen über das Prozesswissen der Schülerinnen und Schüler vor. Ein tiefergehendes Verständnis könnte zu passgenaueren Veränderungen in der Unterrichtsgestaltung führen.

Eine mit dieser Arbeit noch nicht geklärte Fragestellung ist die Beschaffenheit der Teilkompetenz Prozesswissen in der Biologie. Es bleibt zu klären, ob das Prozesswissen eher zu den Sekundärprozessen des prozeduralen Wissens oder zum Problemlösen (Dörner, 1976) zuzuordnen ist. Hierzu sind weitere Validierungsstudien nötig, in denen Tests zu allgemeinem Problemlösen und zu Sekundärprozessen des prozeduralen Wissens simultan zum Prozesswissen in Biologie getestet werden. Wenn folglich Klarheit darüber besteht, welche allgemeinen Kompetenzen mit dem Prozesswissen in Biologie zusammenhängen, können Lernaufgaben und Interventionen zielgerichtet aufgebaut werden.

6.2 Forschungshypothesen B1 sowie B2a und B2b: Zusammenhang von Kompetenz in Biologie mit verwandten kognitiven Konstrukten

Im Zusammenhang mit den Forschungshypothesen B1 sowie B2a und B2b wurde das Strukturgleichungsmodell B aufgestellt. Anhand dieses Modells konnten die Zusammenhänge der übergeordneten Konstrukte sprachliche Kompetenzen und kognitive Grundfähigkeit mit den beiden Teilkompetenzen Konzept- und Prozesswissen von Biologie aufgezeigt werden. Für die Erstellung von Kompetenzentwicklungsmodellen sind die Zusammenhänge und Unterschiede zwischen disziplinspezifischen und -übergreifenden Aspekten von Kompetenz von Bedeutung (Hammann et al., 2007). Die Forschungshypothesen B1 und B2a konnten in ihrer Gänze bestätigt werden; die Hypothese B2b musste aufgrund der Ergebnisse abgelehnt werden. Sprachliche Kompetenzen und

kognitive Grundfähigkeit von Schülerinnen und Schülern leisten einen Beitrag zur Erklärung der beiden Teilkompetenzen, erklären diese in den Regressionen jedoch nicht vollständig. Sprachliche Kompetenzen hängen zudem höher mit Konzeptwissen in Biologie zusammen als kognitive Grundfähigkeit. Kognitive Grundfähigkeit hängt nicht höher mit Konzeptwissen zusammen als sprachliche Kompetenzen. Zudem repliziert die Korrelation $r = .51$ zwischen den beiden externen Kriterien den Befund von Lau und Roeser (2002). Die Instrumente scheinen die externen Kriterien somit gut abzubilden.

Bezogen auf die beiden disziplinübergreifenden Fähigkeiten zeigte sich, dass entgegen der Erwartungen die sprachlichen Kompetenzen in beiden Teilkompetenzen eine bedeutende Rolle spielten. Sowohl schriftlich getestetes Konzept- als auch schriftlich getestetes Prozesswissen werden also stark von den sprachlichen Fähigkeiten der Schülerinnen und Schüler beeinflusst. Die Hinzunahme der beiden disziplinübergreifenden Konstrukte konnte einer weiteren Abgrenzung der beiden getesteten Konstrukte Konzept- und Prozesswissen nur teilweise dienen. Vielmehr bestätigten sie erneut, dass die beiden Teilkompetenzen eng miteinander verwoben sind.

Dieser Befund könnte erneut auf den Methodeneffekt hindeuten. Die vergleichende Betrachtung der Zusammenhänge von sprachlichen Kompetenzen und kognitiver Grundfähigkeit zeigen in der Tendenz, dass letztere für das Prozesswissen von Biologie eine gleichbedeutende Rolle zu spielen scheint wie die sprachlichen Kompetenzen. Hingegen zeigen die sprachlichen Kompetenzen einen höheren Zusammenhang mit dem Konzeptwissen als die kognitive Grundfähigkeit. Wie auch schon der hohe Zusammenhang von Konzept- und Prozesswissen, könnten diese differentiellen Unterschiede aufgrund des Methodeneffekts entstanden sein. Auch diese Schlussfolgerung könnte mit dem bereits vorgeschlagenen simultanen Einsatz von Alternativmethoden und einem schriftlichen Test näher geprüft werden. So liegen zwar zahlreiche korrelative Studien zu den Zusammenhängen zwischen diesen unterschiedlichen Herangehensweisen in der Testung von naturwissenschaftlicher Kompetenz vor (Roberts & Gott, 2004, Shavelson et. al., 2008; Tamir & Doran, 1992; Tamir, Doran & Chye 1992). Zusammenhänge der externen Kriterien im Sinne einer konvergenten und diskriminanten Validitätsprüfung mit den verschiedenen Herangehensweisen fehlen jedoch bisher. Da die unterschiedlichen Testinstrumente jeweils auch unterschiedliche Kompetenzaspekte berücksichtigen, könnte gerade diese externe Validierung Aufschluss darüber geben, wie die Teilaspekte von Kompetenz in Biologie und deren Testungen beschaffen sind.

Sprachliche Kompetenzen tragen also zur Aufklärung der Varianz in beiden Teilkompetenzen von Biologie bei; im Konzeptwissen spielt die Lesekompetenz eine größere Rolle als die kognitive Grundfähigkeit der Schülerinnen und Schüler. Für das Prozesswissen sind beide übergeordneten Konstrukte gleichbedeutend. Im naturwissenschaftlichen Unterricht wird das Prozesswissen neben

Demonstrations- und eigenständig durchgeführten Experimenten sowie weiteren naturwissenschaftlichen Methoden zu einem nicht unbedeutenden Anteil über schriftliche Dokumente wie Tafelbilder, Textbücher oder Arbeitsblätter sowie eigenes Schreiben der Schülerinnen und Schüler und gesprochene Sprache vermittelt (Wellington & Osborne, 2001). Auch diese didaktischen Herangehensweisen könnten zu den hohen Zusammenhängen mit den sprachlichen Kompetenzen für beide Teilbereiche beitragen. Dann läge kein Methodeneffekt vor, sondern ein tatsächlich bestehendes Ineinandergreifen dieser disziplinübergreifenden Kompetenzen mit der disziplinbezogenen Kompetenz. Diese Schlussfolgerung wird durch den differentiellen Effekt der sprachlichen Kompetenz gestützt. Konzeptwissen wird den Schülerinnen und Schülern im unterrichtlichen Geschehen verstärkt über sprachliche Mittel näher gebracht. Im Prozesswissen spielen wiederum die naturwissenschaftlichen Methoden eine größere Rolle. Die kognitive Grundfähigkeit benötigen die Schülerinnen und Schüler in beiden Teilbereichen, um Wissen kognitiv zu verarbeiten und abzuspeichern. Wenn Konzeptwissen stärker über sprachliche Zugänge erlangt wird als Prozesswissen, steht ersteres auch in einem höheren Zusammenhang mit den sprachlichen Kompetenzen der Schülerinnen und Schüler.

Da sprachliche Kompetenzen und kognitive Grundfähigkeit die Kompetenz in Biologie nicht vollständig aufklären können, bleibt die Frage, welche weiteren Aspekte die Kompetenz in Biologie ausmachen. Der eingesetzte Test zu den sprachlichen Kompetenzen besteht aus narrativen Texten mit nicht-naturwissenschaftlichen Inhalten. Eine interessante Folgeuntersuchung könnte naturwissenschaftliche und biologische Texte in das Untersuchungsdesign mit einbeziehen. Da diese eine strenge formale Sprache aufweisen (Wellington & Osborne, 2001; Westby & Torres-Velásquez, 2000), ist zu erwarten, dass diese Texte höher mit der Kompetenz in naturwissenschaftlichen Disziplinen zusammenhängen. Ein weiteres Instrument zur Testung des naturwissenschaftlichen Faktenwissens bzw. des Faktenwissens in der Biologie könnte die linguistischen Aspekte naturwissenschaftlicher Kompetenzen bzw. von Kompetenzen in Biologie komplettieren. Um die kognitiven Grundfähigkeiten um Maße zur naturwissenschaftlichen Kompetenz zu erweitern, würde zum einen ein allgemeiner Problemlösetest, zum anderen ein Test zu naturwissenschaftsbezogenen Problemlösekompetenzen nähere Aufschlüsse dazu geben, welche disziplinübergreifenden Kompetenzen, welche naturwissenschaftliche Kompetenzen und welche Kompetenzen in Biologie Schülerinnen und Schüler benötigen, um Konzept- und Prozesswissen in Biologie zu entwickeln.

Für den Unterricht in Biologie geben die Ergebnisse ebenfalls einige dienliche Hinweise für Lehrerinnen und Lehrer. So zeigen die Resultate erneut, dass Schülerinnen und Schüler gewisse übergreifende Grundvoraussetzungen aus anderen Schulfächern in den Fachunterricht mitbringen müssen, um erfolgreich am unterrichtlichen Geschehen mitwirken zu können. Der hohe

Zusammenhang mit den sprachlichen Kompetenzen zeigt, dass im naturwissenschaftlichen Unterricht Kompetenzen benötigt werden, die die Schülerinnen und Schüler unter anderem im Deutschunterricht erwerben.

Die Erweiterung der Auffassung von Kompetenz in den Naturwissenschaften zu *scientific literacy* wird von fachdidaktischen Forscherinnen und Forschern als auch Bildungswissenschaftlerinnen und Bildungswissenschaftlern gefordert (z. B.: DeBoer, 2000; Bybee, 1997; Gäber & Bolte, 1997). Diese Erweiterung geht mit einer höheren Bedeutung der sprachlichen Kompetenzen einher. Um *scientific literate* zu sein, müssen die Schülerinnen und Schüler die formale Sprache der Naturwissenschaften verstehen, beherrschen und anwenden können. Lehrkräfte der naturwissenschaftlichen Fächer sollten diese eigens naturwissenschaftliche Sprache bei den Schülerinnen und Schülern fördern, um naturwissenschaftliche Kompetenz bei ihnen aufbauen zu können. Der hohe Zusammenhang sollte bei der Unterrichtsgestaltung in den Lehrmaterialien (Fischer et al., 2005) und bei den Lehrkräften also reflektiert werden. Die differentiellen Befunde für das Konzeptwissen zeigen, dass eine höhere sprachliche Kompetenz einer Schülerin bzw. eines Schülers stärker zu einem besseren Konzeptverständnis führt als eine höhere kognitive Grundfähigkeit. Diese eventuell sogar reziproken Zusammenhänge im Entwicklungsprozess könnten längsschnittliche Untersuchungen aufdecken.

Die kognitive Grundfähigkeit konnte in der Vergangenheit als wichtiger Prädiktor für schulische Leistung identifiziert werden (Weinert & Helmke, 1997). Dieser hohe Zusammenhang wurde mit dieser Studie auch für Kompetenz in Biologie bestätigt. Die kognitive Grundfähigkeit sollte daher bei den Schülerinnen und Schülern disziplinübergreifend gefördert werden. Diese Förderung kann in mehreren Schulfächern ermöglicht werden und müsste nach den vorliegenden Erkenntnissen zu Verbesserungen in beiden Teilkompetenzen von Biologie führen. Auch für diese Vermutung könnten die bereits erwähnten Längsschnittstudien nähere Erkenntnisse bringen. Anknüpfungspunkte könnten z. B. in Physik oder Mathematik aber auch im Bereich der Erkenntnisgewinnung in Biologie liegen. Um einen kontinuierlichen Aufbau der kognitiven Grundfähigkeit zu ermöglichen sollten die betreffenden Disziplinen bei der Förderung dieser Fähigkeit disziplinübergreifend zusammenarbeiten.

Die Hypothesentestung war dezidiert jeweils auf die Effekte der übergreifenden Konstrukte jeweils auf das Konzept- und Prozesswissen gerichtet. Es wäre auch möglich, den Blick auf jeweils ein übergreifendes Konzept und dessen differentiellen Effekte auf die beiden Dimensionen von Kompetenz in Biologie zu richten. Damit würde man die Paare von Regressionspfaden vergleichen, die für die Hypothesentestung keine Rolle gespielt haben (kognitive Grundfähigkeit und Konzept- bzw. Prozesswissen sowie sprachliche Kompetenzen und Konzept- bzw. Prozesswissen). So hängt die kognitive Grundfähigkeit vermutlich höher mit dem Prozesswissen zusammen, während für die

sprachlichen Kompetenzen das entgegengesetzte Muster auftreten könnte. Diese würden dann höher mit dem Konzeptwissen zusammenhängen. Diese ersten Hinweise weisen erneut darauf hin, dass mit dem KFT und dem Lückentest valide Instrumente für die beiden übergeordneten Konstrukte anzunehmen sind.

6.3 Fragestellung C: Zusammenhang von Kompetenz in Biologie mit Fachnoten in den naturwissenschaftlichen und nicht-naturwissenschaftlichen Fächern

Mit der Fragestellung C wurden die Fachnoten in das nomologische Netz der Validierung von Kompetenz in Biologie mit seinen Teildimensionen Konzept- und Prozesswissen aufgenommen. Es wurde diskutiert, dass die Notengebung methodische Probleme in sich birgt; ihre praktische Relevanz ist jedoch allgegenwärtig für Schülerinnen und Schüler. So entscheiden die Urteile der sie unterrichtenden Lehrerinnen und Lehrer über den Übertritt in die weiterführenden Schulen oder das Bestehen des Abiturs und somit über die Bildungs- und Berufslaufbahn (Trautwein & Baeriswyl, 2007). Die drei aufgestellten Hypothesen zu den Fachnoten mussten aufgrund der Resultate abgelehnt werden. Die Regressionsgewichte waren für alle Fachnoten in beiden Teildimensionen sehr niedrig und lagen nah beieinander. Es konnten keine differentiellen Regressionsmuster innerhalb der naturwissenschaftlichen Disziplinen und im Vergleich der naturwissenschaftlichen mit den nicht-naturwissenschaftlichen Disziplinen gefunden werden. Dieser Befund stützt die Vermutungen, dass in die Fachnoten, welche stark auf einem gemeinsamen Faktor laden, stark von den gleichen motivationalen und sozialen Einflüssen determiniert werden (Gustafsson & Balke, 1993). Die mittleren Korrelationen vorangegangener Studien zu Zusammenhängen von Fachnoten mit Kompetenzleistungen konnten in den deskriptiven Ergebnissen zwar repliziert werden (Hülür et al., 2011; Schütte et al., 2007; Winkelmann, 2009), in den Regressionen konnten diese Zusammenhänge nicht gezeigt werden. So zeigte die Regression der Fachnote Deutsch auf beide Dimensionen wider Erwarten den höchsten Wert, dicht gefolgt von der Biologie- und Physiknote. Die Physiknote zeigte im Vergleich zur Chemienote einen relativ hohen Zusammenhang in beiden Kompetenzteilbereichen. Hier könnte der von Chiu (2008) erwähnte mathematisch-logische Anteil in der akademischen Disziplin Biologie eine Rolle spielen. Um diese Vermutung zu stützen, müssten weiterführende Analysen durchgeführt werden, die z. B. die kognitive Grundfähigkeit und die Problemlösefähigkeit in Erklärungsmodelle mit einbeziehen. Der bereits mit der Forschungshypothese B aufgedeckte hohe Zusammenhang von Kompetenz in Biologie mit den sprachlichen Kompetenzen der Schülerinnen und Schüler wird durch die Note im Fach Deutsch weiter untermauert. Für das Prozesswissen liegt im Vergleich der naturwissenschaftlichen Fachnoten untereinander zudem ein hoher Zusammenhang mit der Physiknote vor. Dieser könnte auf einen

höheren Anteil an kognitiver Grundfähigkeit hindeuten. Allerdings konnten die bereits berichteten Ergebnisse zur kognitiven Grundfähigkeit diese Vermutung nicht weiter unterstützen.

Das wenig ausdefinierte Muster bestätigt, dass in die Fachnoten weitere Aspekte und nicht nur die Kompetenz in einem Schulfach mit einfließen (Ingenkamp, 1981). Auch mit dieser Untersuchung konnte somit gezeigt werden, dass mit der Beurteilung durch Noten erhebliche Urteilsverzerrungen entstehen könnten (Ingenkamp, 1981). Die Fachnoten einer Schülerin bzw. eines Schülers bestehen nicht nur aus schriftlichen sondern auch aus mündlichen Leistungen und Beobachtungen der Lehrkraft. In die schriftliche Bewertung der Lehrerin bzw. des Lehrers fließen auch die vorangegangenen Leistungen der Schülerin bzw. des Schülers. In die mündlichen Leistungen spielen psychologische Mechanismen wie Ängstlichkeit, Motivation und Leistungsbereitschaft ein. Weitere soziologische Aspekte könnten der sozioökonomische Hintergrund oder das Verhalten in der Gruppe sein. Die Beobachtungen der Lehrkraft wiederum beziehen auch die Persönlichkeit mit ein, also z. B. Fleiß, Aufmerksamkeit oder Pünktlichkeit. Zudem konnte eine vorangegangene Studie zeigen, dass der *big-fish-little-pond-Effekt* (BFLPE) (Marsh, 1987) zu einem bedeutenden Anteil über die vergebenen Schulnoten erklärt werden kann (Trautwein & Baeriswyl, 2007). All diese Merkmale werden durch das stark standardisierte Verfahren eines Kompetenztests soweit wie möglich abgegrenzt. Merkmale, die nicht abgegrenzt werden können, können zumindest kontrolliert werden (z. B. das Selbstkonzept). Daher scheinen Fachnoten für eine Validierung von Kompetenztests nicht sonderlich geeignet.

Ein Grund für die Ablehnung der drei Forschungshypothesen zu den Fachnoten könnten die unterschiedlichen Bezugsnormen sein, die Lehrkräfte bei der Bewertung ihrer Schülerinnen und Schüler anlegen. Die starke schulformabhängige Bezugsnorm der Lehrerinnen und Lehrer bei der Benotung sein. Zum einen wird soziale Bezugsnorm auf zwei Ebenen angelegt. Die Lehrkraft vergibt die Fachnoten innerhalb der sozialen Gruppe einer Klasse. Die Zensuren werden von ihr intuitiv an dem Mittelwert dieser Klasse genormt. Zum anderen haben Lehrerinnen und Lehrer unterschiedlicher Schulformen implizit auch unterschiedliche Erwartungen an ihre Schülerinnen und Schüler. Auch wenn diese Überlegungen nicht überraschend erscheinen mögen, stehen sie im krassen Widerspruch zu ihrer Funktion. So bedeutet die Fachnote zwei in der Gesamtschule nicht das Gleiche wie die Fachnote zwei in einem Gymnasium. Es ist anzunehmen, dass die Lehrkräfte eine allgemeine Einschätzung der Fähigkeiten vornehmen, die sie von Schülerinnen und Schülern der Schulform erwarten, in der sie unterrichten. Diese Einschätzung dient dann als Referenzrahmen für die Benotung einer einzelnen Schülerin bzw. eines einzelnen Schülers in ihrer Klasse. Daher wurden zwei getrennte Modelle gerechnet. Ein Modell mit den Schülerinnen und Schülern, die das Gymnasium besuchen ($n = 1\,496$) und ein Modell mit den Schülerinnen und Schülern, die

Schulformen mit nicht-gymnasialen Bildungsgängen⁷ ($n = 1\,669$) besuchen. Die beiden Modelle führten nicht zu einem differenzierteren Regressionsmuster als im Gesamtmodell (s. Tabelle 39 im Anhang). Aufgrund der offenkundig stark subjektiv determinierten Fachnoten werden tiefgreifende Entscheidungen bezüglich der beruflichen Laufbahn der Schülerinnen und Schüler getroffen (Ingenkamp, 1981). Lehrerinnen und Lehrer sollten für diese Diskrepanz sensibilisiert werden, sodass sie den Faktor der zwischenmenschlichen Beziehungen besser reflektieren können. Zudem wäre eine Ergänzung der subjektiven Einschätzungen der Lehrkräfte durch objektive Verfahren denkbar. Diese müssten aber zweckbezogen auf (prognostische) Validität hin überprüft werden. Mit der Einbeziehung von vorangegangenen Leistungen der Schülerinnen und Schüler ziehen die Lehrkräfte noch eine dritte intraindividuelle Bezugsnorm mit ein.

Ein weiterer Grund für die sehr niedrigen Regressionsgewichte könnte die Grundlage für den vorliegenden Test sein. So war es nicht das Ziel, die Testung inhaltlich möglichst nah an den Curricula der Bundesländer auszurichten. Für einen bundesweiten Test mit naturwissenschaftlichen Inhalten dürfte dies unmöglich sein. Grundlage für den Test waren die Bildungsstandards in Biologie und deren Teilbereiche Fachwissen und Erkenntnisgewinnung. Die Bundesländer sind zwar verpflichtet, diese in ihren Schulen zu implementieren. Dass diese Implementierung noch nicht flächendeckend erreicht ist (Pant et al., 2008), könnte ein weiterer Erklärungsansatz für die niedrigen Zusammenhänge sein. Diese Rechtfertigung der geringen Regressionskoeffizienten sollte jedoch nicht darüber hinweg täuschen, dass womöglich politisch sowohl eine höhere Varianzaufklärung als auch ein höherer Zusammenhang gewünscht sein müsste. Eine gewisse curriculare Validität wäre also auch für diesen Test wünschenswert. Auch wenn subjektive Kriterien in die Bewertung durch die Lehrerinnen und Lehrer eingehen, so sollten standardisierte Tests und Noten besser harmonisieren, als dies im vorliegenden Fall gezeigt werden konnte.

Die Ausführungen zeigen, dass Fachnoten nur bedingt für die diskriminante und konvergente Validierung von Kompetenztests geeignet sind (Lintorf, 2012). Das überraschend niedrige Ergebnismuster führt zu der Frage, was schulische Noten vorgeben zu messen und tatsächlich beinhalten. Somit wäre eine Validierung der Fachnoten – gerade in Bezug auf deren Konsequenzen für die Schülerinnen und Schüler – eine interessante konsekutive Forschungsfrage, die in Deutschland noch nicht so stark im Fokus steht wie z. B. im angelsächsischen Raum (Ingenkamp, 1981). Einschränkend muss jedoch gesagt werden, dass der Bezugsrahmen Klasse und Schulform nur durch die Anpassung der Standardfehler und die Zentrierung auf Klassenebene beachtet wurde. Es ist möglich, dass das Ergebnis zu einem gewissen Teil durch statistische Ungenauigkeit hervorgerufen

⁷ Als Schulformen mit nicht-gymnasialen Bildungsgängen werden im Folgenden Realschulen, Gesamtschulen und Schulen mit mehreren Bildungsgängen zusammengefasst.

wurde. Die getrennten Analysen für die Schulformen deuten darauf hin, dass diese Ungenauigkeit nicht sehr hoch ausfallen dürfte.

6.4 Fragestellung D – Zusammenhang von Kompetenz in Biologie mit naturwissenschaftlichen Selbstkonzepten

Im letzten Schritt wurde das Selbstkonzept als ein wichtiges und bereits gut erforschtes Persönlichkeitsmerkmal zur Bestimmung der Validität des Kompetenztests für Konzept- und Prozesswissen in Biologie aufgenommen. Die bereits vorliegenden Untersuchungen beziehen sich hauptsächlich auf die akademischen Disziplinen Mathematik und Deutsch (Jansen et al, in Revision). Mit dem eingesetzten Instrument zur Messung des Selbstkonzepts ist es gelungen, Selbstkonzept für die drei akademischen Disziplinen Biologie, Physik und Chemie abzubilden. Durch die Differenzierung wurde ersichtlich, dass Selbstkonzept in Biologie unter den naturwissenschaftlichen Disziplinen am meisten zur Aufklärung der Varianz von Kompetenz in Biologie in beiden Teilkompetenzen beiträgt und am höchsten mit beiden Teilkompetenzen zusammenhängt. Die Forschungshypothese D konnte demnach bestätigt werden.

Das vorliegende Ergebnis fügt sich in den bereits bestehenden, ausbaufähigen Forschungsstand zum Selbstkonzept in den naturwissenschaftlichen Disziplinen ein (Jansen et al., in Revision). Selbstkonzepte in den naturwissenschaftlichen Disziplinen scheinen sich für die Biologie, Physik und Chemie getrennt voneinander auszubilden. Dies trifft nicht nur für das schulnähere Konzeptwissen sondern auch für das eher anwendungsbezogene Prozesswissen zu. Hier finden sich sehr ähnliche Zusammenhangsmuster. Im Biologieunterricht scheinen die Schülerinnen und Schüler also nicht ein allgemeines naturwissenschaftliches Selbstkonzept zu bilden, vielmehr wird das spezifische Selbstkonzept in Biologie gefördert. Gleiches könnte auch für die Disziplinen Chemie und Physik gelten. Der Zusammenhang für diese beiden Disziplinen müsste jedoch in Folgeuntersuchungen bestätigt werden.

Da das Selbstkonzept in Biologie in beiden Teildimensionen ungefähr den gleichen Einfluss zeigte, liegt die Vermutung nahe, dass die Förderung des Selbstkonzepts nicht nur positiv auf die Kompetenzen im Konzeptwissen der Schülerinnen und Schüler wirken könnte. Die Ausführungen zu *scientific literacy* haben deutlich gemacht, dass es kaum möglich ist, Schülerinnen und Schüler mit dem notwendigen Wissen für ihr gesamtes Leben auszustatten (Fischer et al, 2005). Das Ziel naturwissenschaftlicher Kompetenzbildung wurde daher erweitert; auch Prozesswissen wird als ein Bestandteil dieser Ausbildung begriffen. Wenn Prozesswissen die Schülerinnen und Schüler nun besser auf ein lebenslanges Lernen im naturwissenschaftlichen Rahmen vorbereiten sollte, spielt das

Selbstkonzept der Schülerinnen und Schüler bei diesem Aspekt naturwissenschaftlicher Bildung eine gleichbedeutende Rolle. Inwiefern die Selbstkonzepte in den naturwissenschaftlichen Disziplinen reziprok entwickelt werden (Marsh, 2002), könnte nur in längsschnittlich angelegten Studien untersucht werden. Weitere Grenzen werden in der übergreifenden Diskussion aufgegriffen.

Das akademische Selbstkonzept wurde als eigenes Bildungsziel expliziert (Marsh & Craven, 2006). Bezüglich des Befundmusters in den Naturwissenschaften stellt sich natürlich die Frage, welches Selbstkonzept dieses Bildungsziel sein soll: das naturwissenschaftliche Selbstkonzept oder das Selbstkonzept in Biologie. Erst wenn diese Frage beantwortet ist, können Förderstrategien beruhend auf empirischen Forschungsergebnissen implementiert werden.

Die gefundenen Zusammenhänge stehen im Einklang mit dem I/E-Modell, welches bisher vorrangig für die Disziplinen Mathematik und Deutsch bestätigt (Lüdtke et al., 2002; Marsh & Köller, 2004) und in neueren Studien auf die Naturwissenschaften erweitert wurde (Chiu 2008, 2012; Jansen et al., in Revision). Das Ergebnis gibt somit darüber hinaus Aufschlüsse darüber, wie sich das Selbstkonzept in den naturwissenschaftlichen Disziplinen entwickeln könnte. Das Selbstkonzept in Biologie steht in Zusammenhang mit den Kompetenzen in Biologie. Die Regressionen auf die weiteren Selbstkonzepte liegen nahe Null. Die Schülerinnen und Schüler scheinen also dezidiert nur die Leistungen innerhalb eines naturwissenschaftlichen Faches zur Formierung ihres Selbstkonzepts zu Rat zu ziehen. Auch für die naturwissenschaftlichen Disziplinen scheinen die Schülerinnen und Schülern differentielle Selbstkonzepte zu entwickeln.

Zusätzlich gibt dieser Befund einen ersten Hinweis darauf, dass das REM-Modell auch in den naturwissenschaftlichen Disziplinen greifen könnte. Die Fachnoten in den Naturwissenschaften korrelieren relativ hoch miteinander, die Selbstkonzepte in den Naturwissenschaften etwas geringer. Das Selbstkonzept in Biologie steht im Zusammenhang mit Kompetenz in Biologie und nicht mit Kompetenz in Chemie und Biologie. Demnach wäre es möglich, dass Selbstkonzept in Biologie auf Kompetenz in Biologie wirkt und *vice versa* während diese gegenseitige Wechselwirkung für Selbstkonzept in Biologie und Kompetenz in Chemie und Physik nicht existiert. Solche reziproken Effekte können jedoch nur anhand von Längsschnittdaten untersucht werden.

Abschließend soll der wichtigste Befund aus den Analysen zur Forschungshypothese D noch einmal herausgestellt werden. Das Selbstkonzept in Biologie steht vorrangig mit Kompetenz in Biologie und nicht mit Kompetenzen in anderen naturwissenschaftlichen Disziplinen im Zusammenhang. Inwiefern dieser Befund ein Ergebnis der getrennten Beschulung naturwissenschaftlicher Disziplinen in Deutschland ist, ist eine interessante weiterführende Fragestellung. Dieser könnte auf Grundlage von internationalen Schulleistungstudien nachgegangen werden, in denen die Schülerinnen und Schüler sowohl zu ihrem Selbstkonzept in den

naturwissenschaftlichen Disziplinen als auch den einzelnen Disziplinen Biologie, Chemie und Physik befragt werden.

6.5 Aspekte der Validierung eines Tests zur Kompetenz in Biologie – Übergreifende Diskussion, Grenzen und Ausblick

Mit der vorliegenden Arbeit wurde zunächst der theoretische Status des Konstrukts Kompetenz in Biologie mit seinen Teildimensionen Konzept- und Prozesswissen näher untersucht. Es wurde versucht zu klären, ob der angewendete Kompetenztest dieses Konstrukt gut abbildet. Der Kompetenztest in Biologie ermöglicht es, zwei Teildimensionen von Kompetenz in Biologie zu testen. Diese werden zu einem hohen Anteil durch sprachliche Kompetenzen und kognitive Grundfähigkeit bestimmt. Differentielle Effekte bezüglich der beiden Teilkompetenzen konnten für die sprachlichen Kompetenzen gefunden werden. Die vermuteten Zusammenhänge mit den Fachnoten unterschiedlicher akademischer Disziplinen konnten nicht bestätigt werden. Die Einbeziehung der Selbstkonzepte in den drei naturwissenschaftlichen Disziplinen konnten zeigen, dass das Selbstkonzept spezifisch für ein naturwissenschaftliches Fach entwickelt werden könnte. Diese zusammengefassten Ergebnisse unterliegen einigen Limitationen, die im Folgenden skizziert werden.

Der Test bestand aus Testaufgaben, die die gleiche Struktur aufwiesen. Nach einem biologischen Einführungstext folgten mehrere Testaufgaben, die von der Schülerin bzw. dem Schüler konsekutiv bearbeitet wurden. Eine der Grundannahmen der IRT ist allerdings, dass die Testaufgaben, die einer Modellierung auf einer Skala zu Grunde liegen, lokal stochastisch voneinander unabhängig sein müssen. Dies bedeutet, dass innerhalb eines Testheftes, das von den Schülerinnen und Schüler bearbeitet wird, die Lösung einer Testaufgabe keine Hinweise für die Lösung einer weiteren Testaufgabe geben darf. Ob diese Vorannahme für den vorliegenden Test zutreffend ist, wurde nur per Augenschein getestet. Weitere analytische Herangehensweisen können näheren Aufschluss zu Abhängigkeiten innerhalb des Tests geben.

Bezogen auf die Validität des Tests konnte mit dieser Arbeit nicht auf alle Aspekte eingegangen werden. So wurden die intendierten Konsequenzen der Testung von Kompetenz in Biologie, und die nicht-intendierten Konsequenzen nicht näher bestimmt und untersucht. Nach dem vereinten Konzept von Validität ist dies jedoch ein immanenter, integraler Bestandteil von Validität. Hierzu wären z. B. detaillierte Abgleiche der Lehrpläne der 16 Bundesländer, Vermutungen zu Einflüssen auf die Unterrichtsgestaltung und die Abschlussprüfungen für den MSA sowie zu intendierten und nicht-intendierten politischen Entscheidungen von Nöten. Einige dieser Aspekte wurden in der Diskussion

aufgenommen; diese ersten Gedanken zur Verwendung des Tests ersetzt jedoch keine systematische Untersuchung dieses Aspektes von Validität.

Mit der vorliegenden Arbeit wurden also nur einige Aspekte der Validität von Kompetenz in Biologie aufgenommen (s. Tabelle 12). So werden nicht die Aufgabenbearbeitungsprozesse und Prozessmodelle untersucht, die hinter der Beantwortung des Tests stehen könnten. Vielmehr wurde eine Konstruktvalidierung durch die Untersuchung der internen Struktur und durch das Inbeziehungsetzen mit externen Kriterien angestrebt. Eine anwendungsbasierte Validierung, auch auf Grundlage der hier referierten Ergebnisse, sollte die Validierung des theoretischen Konstrukts und seiner Testung vervollständigen. Die intraindividuellen Grundlagen für die Performanz im Zusammenhang mit Kompetenz in Biologie würden durch diese Erweiterung näher spezifiziert werden können. Es wurden also – wie dem Titel der Arbeit bereits zu entnehmen ist – Teilaspekte der Validität des Tests zu Konzept- und Prozesswissen in Biologie untersucht. Weitere Validierungen – z. B. bezüglich der intendierten Nutzung des Tests – wären wünschenswert. Doch nicht nur weitere Untersuchungen zur Validität sind vorstellbar.

Um die mit der vorliegenden Studie generierten Ergebnisse zu generalisieren, könnte mit Studien in weiteren Altersgruppen (z.B. im Grundschulalter) geprüft werden, ob die Ergebnisse repliziert werden können bzw. anhand von Längsschnittstudien das Verhältnis der beiden Teildisziplinen zueinander im Laufe der Kompetenzentwicklung in den Naturwissenschaften untersucht werden. Hierzu wären auch Experimentalstudien denkbar, wie sie zu den beiden Teildimensionen in der Mathematik durchgeführt wurden (Perry, 1991; Rittle-Johnson et al., 2001; Rittle-Johnson & Wagner Alibali, 1999). Wenn Entwicklungsfortschritte in der Kompetenz in den Naturwissenschaften der Schülerinnen und Schüler in den Teilkompetenzen unterschiedlich verlaufen, würden eindimensionale Modelle – auch im Zusammenhang mit der Berichterstattung in *large-scale assessment* Studien – wichtige Charakteristika der Zuwächse in den einzelnen Dimensionen verschleiern. Dies könnte eine Berichterstattung auf Grundlage von profunden mehrdimensionalen Modellen leisten.

In der vorliegenden Arbeit wurde der Schwerpunkt auf eine naturwissenschaftliche Disziplin gesetzt. Daher konnte nicht geklärt werden, inwiefern die beiden Teildimensionen Konzept- und Prozesswissen originär biologische oder naturwissenschaftliche Teilkompetenzen sind. Erste Hinweise führt die Studie von Senkbeil et al. (2005) an. Es konnten sieben allgemeine kognitive Dimensionen modelliert werden; die inhaltlichen Dimensionen Biologie, Chemie und Physik bildeten keine eigenen Dimensionen. Bezogen auf die in der vorliegenden Arbeit operationalisierte naturwissenschaftsbezogene Prozessdimension wären weitere interne Strukturen denkbar. So könnten drei inhaltliche Dimensionen einer gemeinsamen naturwissenschaftlichen Prozessdimension

gegenüberstehen. In weiterführenden Studien, die alle drei naturwissenschaftlichen Disziplinen umfassen, könnte diese Frage geklärt werden. Da die unterschiedlichen Konzeptualisierungen von *scientific literacy* verschiedene Aspekte in die Prozesskomponente aufnehmen, könnten weitere Aspekte wie Bewertung oder Kommunikation Einblicke in die Beschaffenheit weiterer Prozessdimensionen geben.

Die hinter einem Test stehende Konzeptualisierung von *scientific literacy* hat Folgen bezogen auf dessen intendierte und nicht-intendierte Nutzungen. Wenn – insbesondere in deutschlandweiten Testungen wie dem an PISA angekoppelten Ländervergleich oder den Vergleichsarbeiten – nur ein Teilaspekt, z. B. das Konzeptwissen, getestet wird, könnte dies unerwünschte Folgen in Bezug auf den Unterricht in den naturwissenschaftlichen Fächern (Baartman et al., 2007) oder bildungspolitische Entscheidungen haben. Nationale *large-scale assessment* Studien geben eine Momentaufnahme zu den Ergebnissen des Bildungssystems. Auch eine Testung nur eines Teilaspekts von Prozesswissen könnte z. B. dazu führen, dass nur dieser Teilaspekt zu Lasten anderer Teilaspekte in die Curricula der Bundesländer einfließt. Rückgreifend auf die Erläuterungen zur Kompetenz in Biologie in der Schule (s. Seite 32) entspräche die schriftliche Testung des Prozesswissens innerhalb des Curriculums des Landes Berlin nur dem Punkt des theoretischen Arbeitens und schliesse das Durchführen und Auswerten von Experimenten zu Photosynthese und Atmung aus. Die einseitige Testung würde dann zu einer Verbesserung der schulischen Leistungen in Biologie und einem tieferen Verständnis auf isolierten Teilgebieten führen. Der vernetzte, allumfassende Wissensbestand und insbesondere der Handlungsaspekt der Erkenntnisgewinnung in den Naturwissenschaften würde im unterrichtlichen Geschehen nur eine untergeordnete Rolle spielen. Gerade dieser Aspekt wurde von Bildungsforscherinnen und –forschern in der Vergangenheit jedoch als vernachlässigt und förderungsbedürftig erkannt (Baumert et al., 1997; Duit et al., 2001; Hammann, 2006; Klos et al., 2008; PISA-Konsortium Deutschland, 2007). Eine breite Abdeckung der jeweils getesteten Kompetenzbereiche ist somit unabdingbar. Mit der vorliegenden Arbeit konnte gezeigt werden, dass dies mit einer schriftlichen Testung nicht gewährleistet werden kann. Den Rezipienten von *large-scale assessment* Studien sollte diese Limitation kommuniziert werden, um die Validität gewährleisten zu können. Mit dieser Forderung geht auch einher, dass weitere Kompetenzbereiche, wie die Bewertungs- und Kommunikationskompetenz in deutschlandweite Testungen von Kompetenz in den Naturwissenschaften Eingang finden müssen.

Diese Forderung gilt nicht nur über Kompetenzbereiche hinweg, sondern auch innerhalb eines Kompetenzbereichs. So wurde Prozesswissen im Rahmen von *large-scale assessment* Studien bisher selten eigenständig und differenziert getestet. Aufgrund des angewandten Testdesigns war es mit der vorliegenden Untersuchung nicht möglich, unterschiedliche Facetten innerhalb des

Prozesswissens abzubilden, z. B. das Experiment, die Beobachtung, den Vergleich oder das Arbeiten mit Modellen, wie es Senkbeil et al. (2005) für kognitive Anforderungen bereits gelungen ist. Diese Untersuchungen werden in kleineren Studien – z. B. in der Chemie und Biologie – bereits durchgeführt (Nehring et al., 2011). Analysen in diesen Bereichen sind in zukünftigen Untersuchungen anzustreben, da sie Aufschluss darüber geben, wie das Prozesswissen strukturiert ist und in welchen Teilaspekten die Schwächen und Stärken der Schülerinnen und Schüler liegen. Erste Ergebnisse hierzu konnten Mayer et al. (2008) für die Biologie vorlegen. Eine konsequente Frage ist, ob das Prozesswissen in den drei naturwissenschaftlichen Disziplinen unterschiedlich gestaltet ist und somit drei unterschiedliche Wege der Erkenntnisgewinnung vorliegen oder ob es einen naturwissenschaftlichen Erkenntnisweg gibt. Diese Frage könnte mit der Erhebung von Konzept- und Prozesswissen für alle drei naturwissenschaftlichen Disziplinen und anschließenden Dimensionalitätsanalysen beantwortet werden.

Mehrere Hypothesentestungen in der vorliegenden Arbeit legten die Vermutung nahe, dass durch die schriftliche Testung des Prozesswissens, der gefundene Zusammenhang auf einem Methodeneffekt beruhen könnte. So führte dieser Effekt aller Wahrscheinlichkeit nach zu der sehr hohen Korrelation der beiden Teildimensionen und zu starken Einflüssen der externen Kriterien sprachliche Kompetenzen und Note im Fach Deutsch sowohl auf das Konzept- als auch auf das Prozesswissen. Mit der hier vorgenommenen schriftlichen Testung des Prozesswissens wurde nur eine von vielen Erhebungsmöglichkeiten angewandt. Bei der Durchführung einer Studie mit einem großen Stichprobenumfang ist dies die am besten handhabbare Möglichkeit, um die Kompetenz aller Schülerinnen und Schüler reliabel und valide zu messen. Diese Möglichkeit scheint jedoch stark mit dem Konzeptwissen der Schülerinnen und Schüler zusammen zu hängen und zeigt hohe Zusammenhänge mit den bereits genannten externen Kriterien. Diese Befunde lassen darauf schließen, dass mit der vorliegenden Testung nur ein Teilaspekt von Prozesswissen abgedeckt wurde. Für diesen Teilaspekt konnte gezeigt werden, dass er hoch mit dem Konzeptwissen in Biologie zusammenhängt. Zudem hängen sowohl das Konzept- als auch das hier abgebildete Prozesswissen hoch mit den sprachlichen Kompetenzen und kognitiver Grundfähigkeit zusammen und werden durch das Selbstkonzept in Biologie beeinflusst. Die Zusammenhänge mit den Fachnoten unterschiedlicher akademischer Disziplinen lassen darauf schließen, dass in die Benotung durch Lehrkräfte viele verschiedene Beurteilungsprozesse einfließen. Daher scheinen Noten nicht zur Validierung von Kompetenztests geeignet zu sein. Die bereits erwähnten alternativen Testmethoden sollten – trotz ihrer Unzulänglichkeiten bezogen auf Reliabilität auch im *large-scale* Bereich – in Untersuchungsdesigns aufgenommen werden. Über diese multiple Herangehensweise könnte die jeweilige Testung verortet werden.

7 Anhang

Tabelle 29: Inhaltliche Beschreibung der Dimensionen naturwissenschaftlicher Kompetenz konzeptualisiert bei PISA 2006 (OECD, 2007; PISA-Konsortium Deutschland, 2007)

Wissenskomponente	Prozesskomponente
Kennen naturwissenschaftlicher Fakten, Terme und Konzepte (<i>knowledge of science</i>)	Naturwissenschaftliche Fragestellungen erkennen: <ul style="list-style-type: none"> • Erkennen von Fragestellungen, die mit naturwissenschaftlichen Methoden untersucht werden können • Schlüssel- bzw. Stichwörter für die Suche nach naturwissenschaftlichen Informationen identifizieren • grundlegende Eigenschaften einer naturwissenschaftlichen Untersuchung verstehen
Entwicklung von Verständnis naturwissenschaftlicher Prinzipien, Gesetze, Methoden, Prozesse, Konzepte und Grenzen (<i>knowledge about science</i>)	naturwissenschaftliche Phänomene erklären: <ul style="list-style-type: none"> • vorhandenes Wissen in einer gegebenen Situation anwenden • naturwissenschaftliche Phänomene beschreiben oder interpretieren • Änderungen vorhersagen und angemessene Beschreibungen, Erklärungen und Vorhersagen erkennen
Fokussierung auf thematische Inhalte	Naturwissenschaftliche Evidenz nutzen: <ul style="list-style-type: none"> • naturwissenschaftliche Beweise interpretieren und gültige Schlussfolgerungen ziehen und kommunizieren • Annahmen, Beweise und Argumentationen erkennen, die zu Schlussfolgerungen führen • über die gesellschaftliche Bedeutung von Fortschritten in Naturwissenschaften und Technik reflektieren können • Fähigkeit, angemessene und sichere Schlussfolgerungen aus Beweisen und Informationen ziehen zu können • auf der Basis von Beweisen Behauptungen anderer kritisieren, Meinungen von auf Beweisen basierenden Aussagen unterscheiden
	Naturwissenschaftliches Forschen: <ul style="list-style-type: none"> • Ausgangspunkt (z. B. naturwissenschaftliche Fragen) • Ziele (z. B. theoriegeleitete Untersuchungen) • Experimente (z. B. experimentelles Design) • Daten (z. B. quantitative, qualitative) • Messungen (z. B. Replizierbarkeit) • Charakteristische Eigenschaften der Ergebnisse (z. B. vorläufig) <p>Fortsetzung nächste Seite</p>

	<p>Naturwissenschaftliche Erklärungen:</p> <ul style="list-style-type: none"> • Typen (z.B. Hypothesen, Modelle, Theorien) • Entstehung (z. B. Logik) • Regeln (z.B. Erklärungen müssen logisch konsistent sein und auf Beweisen, historischem und aktuellem Wissen beruhen) • Ergebnisse (z. B. Erklärungen führen zu neuen Fragestellungen und Untersuchungen)
--	--

Tabelle 30: Beispielhafte inhaltliche Ausgestaltung der Basiskonzepte des Kompetenzbereichs Fachwissen im Fach Biologie (KMK, 2005a)

Basiskonzept	Inhalt anhand von Beispielen
System	<p>Biosysteme</p> <p>Elemente von lebendigen Systemen</p> <p>Eigenschaften von lebendigen Systemen und Wechselwirkung zwischen belebter und unbelebter Natur</p> <p>Struktur und Funktion lebendiger Systeme</p>
Struktur und Funktion	<p>Erfassen, Ordnen und Wiedererkennen von Strukturen</p> <p>Strukturelle und funktionelle Grundbaueinheit des Lebendigen – die Zelle</p> <p>Funktionen von Zelle, Organismen, Ökosystem und Biosphäre</p>
Entwicklung	<p>Unterscheidung von Individual- und evolutionärer Entwicklung</p> <p>Artspezifische Entwicklung von Zelle und Organismus</p> <p>Genetische Anlagen und Umwelteinflüsse</p>

Tabelle 31: Inhaltliche Ausformulierung des Kompetenzbereichs Erkenntnisgewinnung im Fach Biologie (KMK, 2005a)

Methode	Inhalt
Kriteriengeleitete Beobachtung	<p>Beinhaltet drei Schritte:</p> <ul style="list-style-type: none"> • Formulierung einer Fragestellung
Hypothesengeleitetes Arbeiten	<ul style="list-style-type: none"> • Hypothesenbildung, Planung einer Beobachtung, eines Vergleichs, eines Experiments und Durchführung der Untersuchungsmethode • Werten und Interpretieren der gewonnenen Daten
Kriterienbezogenes Vergleichen	<p>Erlangung biologischer Erkenntnisse über verwandtschaftliche Beziehungen</p> <p>Schließt Auseinandersetzung mit der Evolutionstheorie mit ein</p>
Modellbildung	<p>Veranschaulichung komplexer Phänomene</p> <p>Modellieren bzw. kritisches Reflektieren eines Modells</p>

Tabelle 32: Standards für den Kompetenzbereich Fachwissen (KMK, 2005a)

Die Schülerinnen und Schüler ...

F1 System

- F1.1 verstehen die Zelle als System,
- F1.2 erklären den Organismus und Organismengruppen als System,
- F1.3 erklären Ökosystem und Biosphäre als System,
- F1.4 beschreiben und erklären Wechselwirkungen im Organismus, zwischen Organismen sowie zwischen Organismen und unbelebter Natur,
- F1.5 wechseln zwischen den Systemebenen,
- F1.6 stellen einen Stoffkreislauf sowie den Energiekreislauf in einem Ökosystem dar,
- F1.7 beschreiben Wechselwirkungen zwischen Biosphäre und den anderen Sphären der Welt,
- F1.8 kennen und verstehen die grundlegenden Kriterien von nachhaltiger Entwicklung.

F2 Struktur und Funktion

- F2.1 beschreiben Zellen als strukturelle und funktionelle Grundbaueinheiten von Lebewesen,
- F2.2 vergleichen die bakterielle, pflanzliche und tierliche Zelle in Struktur und Funktion,
- F2.3 stellen strukturelle und funktionelle Gemeinsamkeiten und Unterschiede von Organismen und Organismengruppen dar,
- F2.4 beschreiben und erklären Struktur und Funktion von Organen und Organsystemen, z. B. bei der Stoff- und Energieumwandlung, Steuerung und Regelung, Informationsverarbeitung, Vererbung und Reproduktion,
- F2.5 beschreiben die strukturelle und funktionelle Organisation im Ökosystem,
- F2.6 beschreiben und erklären die Anpasstheit ausgewählter Organismen an die Umwelt.

F3 Entwicklung

- F3.1 erläutern die Bedeutung der Zellteilung für Wachstum, Fortpflanzung und Vermehrung,
- F3.2 beschreiben die artspezifische Individualentwicklung von Organismen,
- F3.3 beschreiben verschiedene Formen der Fortpflanzung,
- F3.4 beschreiben ein Ökosystem in zeitlicher Veränderung,
- F3.5 beschreiben und erklären stammesgeschichtliche Verwandtschaft von Organismen,
- F3.6 beschreiben und erklären Verlauf und Ursachen der Evolution an ausgewählten Lebewesen,
- F3.7 erklären die Variabilität von Lebewesen,
- F3.8 kennen und erörtern Eingriffe des Menschen in die Natur und Kriterien für solche Entscheidungen.

Aus *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss* (S. 13-14) von KMK, 2005, München: Wolters Kluwer Deutschland.

Tabelle 33: Standards für den Kompetenzbereich Erkenntnisgewinnung (KMK, 2005a)

Die Schülerinnen und Schüler ...	
E1	mikroskopieren Zellen und stellen sie in einer Zeichnung dar,
E2	beschreiben und vergleichen Anatomie und Morphologie von Organismen,
E3	analysieren die stammesgeschichtliche Verwandtschaft bzw. ökologisch bedingte Ähnlichkeit bei Organismen durch kriteriengeleitetes Vergleichen,
E4	ermitteln mithilfe geeigneter Bestimmungsliteratur im Ökosystem häufig vorkommende Arten,
E5	führen Untersuchungen mit geeigneten qualifizierenden oder quantifizierenden Verfahren durch,
E6	planen einfache Experimente, führen die Experimente durch und/oder werten sie aus,
E7	wenden Schritte aus dem experimentellen Weg der Erkenntnisgewinnung zur Erklärung an,
E8	erörtern Tragweite und Grenzen von Untersuchungsanlage, -schritten und -ergebnissen,
E9	wenden Modelle zur Veranschaulichung von Struktur und Funktion an,
E10	analysieren Wechselwirkungen mit Hilfe von Modellen,
E11	beschreiben Speicherung und Weitergabe genetischer Information auch unter Anwendung geeigneter Modelle,
E12	erklären dynamische Prozesse in Ökosystemen mithilfe von Modellvorstellungen,
E13	beurteilen die Aussagekraft eines Modells.

Aus *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss* (S. 14) von KMK, 2005, München: Wolters Kluwer Deutschland.

Tabelle 34: Mittlere Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für den Kompetenzbereich Fachwissen (SBJS, 2006)

System-Konzept

Die Schülerinnen und Schüler ...

erklären Grundsätze des Nervens- und Hormonsystems als sich ergänzende Systeme der Informationsübermittlung.

Struktur- und Funktions-Konzept

Die Schülerinnen und Schüler ...

stellen Fotosynthese und Zellatmung als Prinzipien der Energieversorgung dar.

erkennen biologische Regeln sowie Gesetzmäßigkeiten und wenden sie unter Anleitung an.

Entwicklungs-Konzept

Die Schülerinnen und Schüler ...

beschreiben die menschliche Entwicklung von der Zygote bis zum Tod.

erläutern die Individualentwicklung von Organismen im Überblick.

unterscheiden verschiedene Formen der Fortpflanzung.

analysieren und erklären stammesgeschichtliche Verwandtschaft von Organismen.

erklären Organsysteme und deren Anpasstheit im Zusammenhang zur Evolutionstheorie.

untersuchen und erklären den Verlauf stammesgeschichtlicher Entwicklung an ausgewählten Lebewesen.

erklären die Variabilität von Lebewesen.

Aus Rahmenlehrplan für die Sekundarstufe 1 (S. 18), SBJS, 2006, Berlin: Oktoberdruck.

Tabelle 35: Einfache, mittlere und erweiterte Standards des Berliner Rahmenlehrplans Biologie für die Doppeljahrgangsstufe 9/10 für den Kompetenzbereich Fachwissen – Entwicklungs-Konzept

Entwicklungs-Konzept		
Einfache Standards	Mittlere Standards	Erweiterte Standards
Die Schülerinnen und Schüler ...		
<ul style="list-style-type: none"> • beschreiben die menschliche Entwicklung von der Zygote bis zum Tod. • stellen die Individualentwicklung eines Organismus dar. • beschreiben eine Form der Fortpflanzung. • beschreiben Organsysteme und deren Angepasstheit im Zusammenhang zur Evolutionstheorie. • untersuchen den Verlauf stammesgeschichtlicher Entwicklung an einem gewählten Lebewesen. 	<ul style="list-style-type: none"> • erläutern die Individualentwicklung von Organismen im Überblick. • unterscheiden verschiedene Formen der Fortpflanzung. • analysieren und erklären stammesgeschichtliche Verwandtschaft von Organismen. • erklären Organsysteme und deren Angepasstheit im Zusammenhang zur Evolutionstheorie. • untersuchen und erklären den Verlauf stammesgeschichtlicher Entwicklung an ausgewählten Lebewesen. • erklären die Variabilität von Lebewesen. 	<ul style="list-style-type: none"> • erklären die menschliche Individualentwicklung unter den Aspekten Zelldifferenzierung, Arbeitsteilung und Altern. • erklären verschiedene Formen der Fortpflanzung.

Anmerkungen. Aus „Rahmenlehrplan für die Sekundarstufe I“, von SBJS, 2006, S. 18. Berlin: Oktoberdruck.

Tabelle 36: Messfehlerbereinigte Korrelationen der Fachnoten Biologie, Chemie, Physik, Deutsch und Mathematik

Fachnote	Deutsch	Mathematik	Biologie	Chemie
Mathematik	.39			
Biologie	.51	.45		
Chemie	.42	.58	.56	
Physik	.43	.60	.52	.60

Tabelle 37: Regressionsgewichte der Fachnoten Chemie, Physik, Mathematik und Deutsch für die beiden Teildimensionen Konzept- und Prozesswissen in Biologie und deren Standardfehler

Fachnote	Konzeptwissen ¹		Prozesswissen ¹	
	β	<i>SE</i>	β	<i>SE</i>
Deutsch	.16*	.03	.16*	.03
Mathematik	.09	.04	.09	.03
Chemie	.05	.03	.04	.03
Physik	.10*	.04	.11*	.04

Anmerkungen. ¹ $R^2 = .10$.

β = Standardisiertes Regressionsgewicht, *SE* = Standardfehler.

* = Regressionskoeffizienten sind auf einem Niveau von $\alpha = .05$ statistisch signifikant von Null verschieden.

Tabelle 38: Regressionsgewichte der Teildimensionen Konzept- und Prozesswissen auf die Selbstkonzepte Chemie und Physik

Selbstkonzept	Konzeptwissen ¹		Prozesswissen ¹	
	β	<i>SE</i>	β	<i>SE</i>
Chemie	.01	.04	.04	.04
Physik	.17*	.04	.15*	.05

Anmerkungen. ¹ $R^2 = .03$.

* = Regressionskoeffizienten sind auf einem Niveau von $\alpha = .05$ signifikant von Null verschieden.

β = Standardisiertes Regressionsgewicht, *SE* = Standardfehler

Tabelle 39: Regressionsgewichte der Teildimensionen Konzept- und Prozesswissen in Biologie auf die Fachnoten Biologie, Chemie, Physik, Mathematik und Deutsch, Teilpopulationen Gymnasien ($n = 1\,496$) und Schulformen mit nicht-gymnasialen Bildungsgängen ($n = 1\,669$)

Fachnote	Gymnasium				Nicht-Gymnasium			
	Konzeptwissen ¹		Prozesswissen ²		Konzeptwissen ³		Prozesswissen ⁴	
	β	SE	β	SE	β	SE	β	SE
Deutsch	.10*	.04	.16*	.05	.13*	.04	.18*	.04
Mathematik	.15	.06	.10	.05	.12	.05	.08*	.05
Biologie	.16*	.05	.12*	.06	.07	.05	.08	.04
Chemie	.00	.05	-.01	.06	.04	.04	.02	.05
Physik	.14*	.06	.19*	.05	.04	.04	.08	.05

Anmerkungen. * = Regressionskoeffizienten sind auf einem Niveau von $\alpha = .05$ signifikant von Null verschieden.

Varianzaufklärung der einzelnen latenten Faktoren in den beiden Modellen: $^1R^2 = .16$, $^2R^2 = .20$, $^3R^2 = .16$, $^4R^2 = .09$.

β = Standardisiertes Regressionsgewicht, SE = Standardfehler.

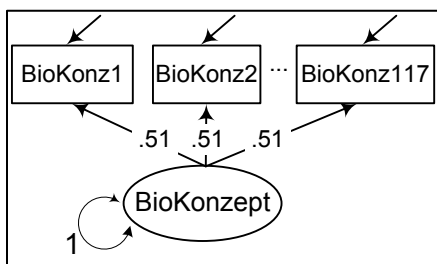


Abbildung 26: Messmodell für Biologie Konzeptwissen

Anmerkungen. Bio = Biologie, Konz = Konzeptwissen.

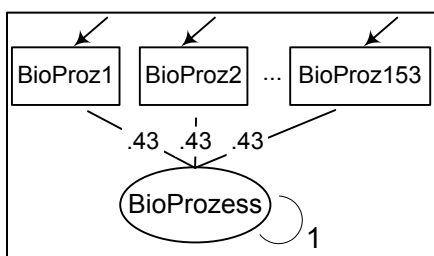


Abbildung 27: Messmodell für Biologie Prozesswissen

Anmerkungen. Bio = Biologie, Proz = Prozesswissen.

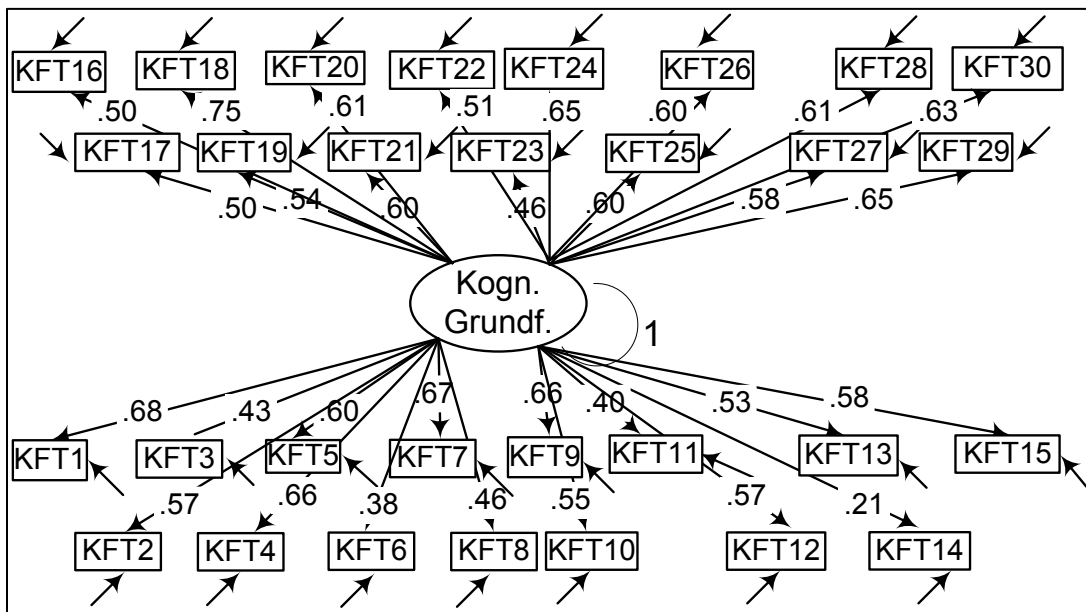


Abbildung 28: Messmodell für kognitive Grundfähigkeit

Anmerkungen. KFT, Kogn. Grundf. = Kognitiver Fähigkeitstest.

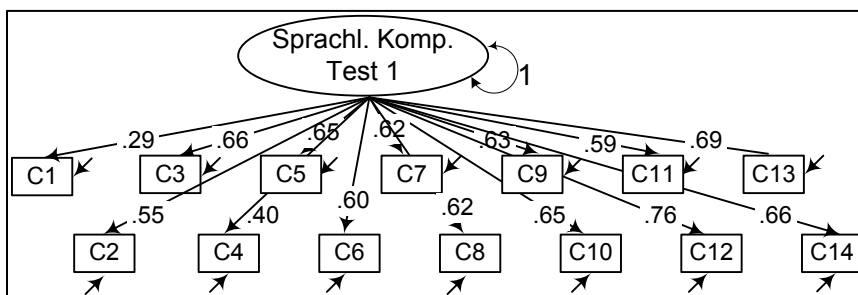


Abbildung 29: Messmodell für sprachliche Kompetenzen – Lückentest 1

Anmerkungen. Sprachl. Komp = sprachliche Kompetenzen, C = Cloze-Test.

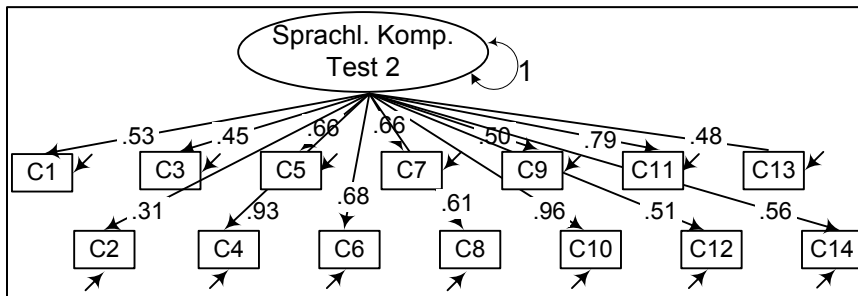


Abbildung 30: Messmodell für sprachliche Kompetenzen – Lückentest 2

Anmerkungen. Sprachl. Komp = sprachliche Kompetenzen, C = Cloze-Test.

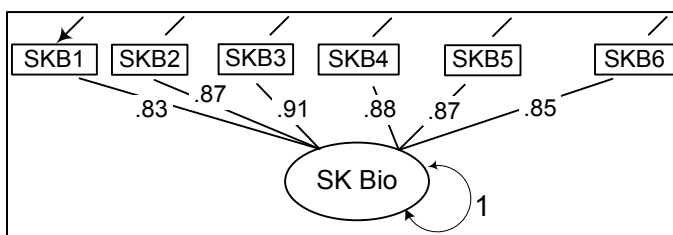


Abbildung 31: Messmodell für Selbstkonzept in Biologie

Anmerkungen. SKB, SK Bio = Biologie Selbstkonzept.

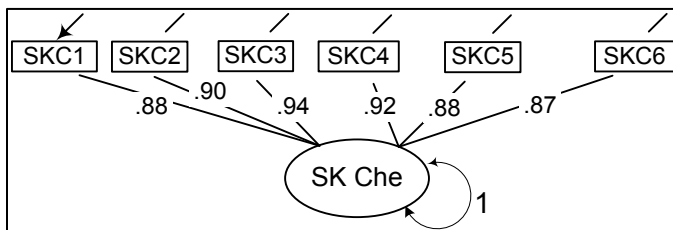


Abbildung 32: Messmodell für Selbstkonzept in Chemie

Anmerkungen. SKC, SK Che = Chemie Selbstkonzept.

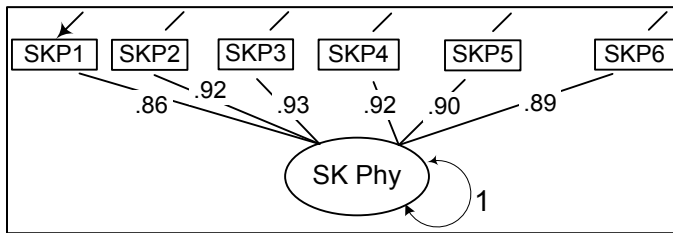


Abbildung 33: Messmodell für Selbstkonzept in Physik

Anmerkungen. SKP, SK Phy = Physik Selbstkonzept.

8 Literaturverzeichnis

- Ackerman, P. L. (1989). Individual differences and skill acquisition. In P. L. Ackerman, R. J. Steinberg & R. Glaser (Hrsg.), *Learning and individual differences: advances in theory and research*. New York: Freeman.
- Ackerman, P. L., Beier, M. E. & Bowen, K. R. (2000). Explorations of crystallized intelligence. Completion tests, cloze tests, and knowledge. *Learning and Individual Differences*, 12, 105-121.
- Adams, R. & Wu, M. (Hrsg.). (2002). *PISA 2000 technical report*. Paris: OECD.
- American Association for the Advancement of Science. (1990). *Science for all americans*. New York: Oxford University Press.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- American Educational Research Association. (2004). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Alexander, P. A. & Judy, F. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, 58, 375-404.
- Allen, N. L., Carlson, J. E. & Zelenak, C. A. (1999). *The NAEP 1996 Technical Report*. Washington, DC: National Center for Statistics.
- Anderson, J. R. (1976). *Language, memory and thought*. Hillsday, NJ: Erlbaum.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Anderson, J. R. (1989). *Kognitive Psychologie. Eine Einführung*. New York: W.H. Freeman and Company.
- Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum Akademischer Verlag.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., . . . Wittrock, M. C. (Hrsg.). (2001). *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

- Arbinger, R. (1998). Komplexität bei der Entwicklung und dem Aufbau von Wissensstrukturen. *Geographie und Schule*, 20 (116), 25-32.
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A. & van der Vleuten, C. P. M. (2007). Evaluating assessment quality in competence-based education: a qualitative comparison of two frameworks. *Educational Research Review*, 2, 114-129.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden*. Berlin: Springer.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556-559.
- Baumert, J. (2002). Deutschland im internationalen Vergleich. In N. Killius, J. Kluge & L. Reisch (Hrsg.), *Die Zukunft der Bildung* (S. 100-150). Frankfurt am Main: Suhrkamp.
- Baumert, J., Bos, W. & Lehmann, R. (Hrsg.). (2000). *TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie - Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2 Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe*. Opladen: Leske + Budrich.
- Baumert, J., Brunner, M., Lüdtke, O. & Trautwein, U. (2007). Was messen internationale Schulleistungsstudien? - Resultate kumulativer Wissenserwerbsprozesse. Eine Antwort auf Heiner Rindermann. *Psychologische Rundschau*, 58, 118-128.
- Baumert, J., Lehmann, R. & Lehrke, M. (1997). *TIMSS - Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske + Budrich.
- Baumert, J., Lüdtke, O., Trautwein, U. & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: evidence in support of the distinction between intelligence and student achievement. *Educational Research Review*, 4, 165-176.
- Bayrhuber, H. (2007). Entwicklung von Bildungsstandards für den Mittleren Schulabschluss in der Sekundarstufe 1 - Biologie, Chemie, Physik. In H. Bayrhuber, D. Elster, D. Krüger & H. J. Vollmer (Hrsg.), *Kompetenzentwicklung und Assessment. Forschungen zur Fachdidaktik. Band 9* (S. 69-78). Innsbruck: Studienverlag.

- Bayrhuber, H., Bögeholz, S., Eggert, S., Elster, D., Grube, C., Hössle, C., . . . Schoormans, G. (2007). Biologie im Kontext - Erste Forschungsergebnisse. *Der mathematische und naturwissenschaftliche Unterricht*, 60, 304-313.
- Beaton, A. E., Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Smith, T. A. & Kelly, D. L. (1996). *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.
- Beghetto, R. A. (2007). Factors associated with middle and secondary students' perceived science competence. *Journal of Research in Science Teaching*, 44, 800-814.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 170, 238-246.
- Bernholt, S., Parchmann, I. & Commons, M. L. (2009). Kompetenzmodellierung zwischen Forschung und Unterrichtspraxis. *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 219-245.
- Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology*, 36, 209-231.
- Bloom, B. S. (1974). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim und Basel: Beltz.
- Bollen, K. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bond, L. (1989). The effects of special preparation on measures of scholastic ability. In R. L. Linn (Hrsg.), *Educational measurement* (S. 429-444). New York: Macmillan.
- Borsboom, D., Mellenbergh, G. J. & Heerden, J. van (2004). The concept of validity. *Psychological Review*, 111, 1061-1071.
- Brown, B. A., Reveles, J. M. & Kelly, G. J. (2005). Scientific literacy and discursive identity: a theoretical framework for understanding science learning. *Science Education*, 89, 779-802.

- Brunner, M. (2006). Mathematische Schülerleistung: Struktur, Schulformunterschiede und Validität, Zugriff am 6. 10. 2012 <http://edoc.hu-berlin.de/dissertationen/brunner-martin-2006-02-08/HTML/front.html>
- Burt, C. (1949). The structure of mind: a review of the results of factor analysis. *British Journal of Educational Psychology*, 19, 176-199.
- Bybee, R. W. (1997). Toward an understanding of scientific literacy. In W. Gräber & C. Bolte (Hrsg.), *Scientific literacy* (S. 37-68). Kiel: IPN.
- Bybee, R. W. (2002). Scientific literacy. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific literacy. Der Beitrag der naturwissenschaften zur Allgemeinen Bildung* (S. 21-43). Opladen: Leske + Budrich.
- Bybee, R. W., McCrae, B. & Laurie, R. (2009). PISA 2006: an assessment of scientific literacy. *Journal of Research in Science Teaching*, 46(8), 865-883.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Campbell, N. A. & Reece, J. B. (2003). *Biologie*. Heidelberg: Spektrum.
- Carroll, J. B. (1993). *Human cognitive abilities: a survey of factoranalytic studies*. New York: Cambridge University Press.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities: past, present and future. In D. P. Flanagan & P. L. Harrison (Hrsg.), *Contemporary intellectual assessment. Theories, tests and issues* (S. 69-76). New York: Guilford Press.
- Carter, R. S. (1952). How invalid are marks assigned by teachers. *Journal of Educational Psychology*, 43, 218-228.
- Cattell, R. B. (1987). *Intelligence: its structure, growth, and action*. Boston: Houghton Mifflin.
- Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, 38, 592.
- Cattell, R. B. (1943). The measurement of adult education. *Psychological Bulletin*, 40, 153-193.

- Chinn, C. A. & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research*, 63, 1-49.
- Chiu, M.-S. (2008). Achievements and self-concepts in a comparison of math and science: exploring the internal/external frame of reference model across 28 countries. *Educational Research and Evaluation*, 14, 235-254.
- Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-little-pond effect and combined model for mathematics and science. *Journal of Educational Psychology*, 104, 87-107.
- Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-301.
- Cummins, J. (2008). BICS and CALP: empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Hrsg.), *Encyclopedia of language and education* (S. 71-83). New York: Springer Science + Business Media LLC.
- Deary, I. J., Strand, S., Smith, P. & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35, 13-21.
- DeBoer, G. E. (2000). Scientific literacy: another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, 37, 582-601.
- Denig, F. & Weis, V. (1970). Zur Faktorenstruktur der Fachnoten in Sekundärschulen. *Zeitschrift für Empirische Forschung*, 4, 210-232.
- Destatis. (2011). *Bildung und Kultur. Berufliche Bildung*. Wiesbaden: Statistisches Bundesamt.
- Deutsches PISA-Konsortium (Hrsg.). (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Deutsches PISA-Konsortium (Hrsg.). (2002). *PISA 2000 - Die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.

- Deutsches PISA-Konsortium (Hrsg.). (2003). *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen: Leske + Budrich.
- Doran, R. L. & Tamir, P. (1992). Results of practical skills testing. *Studies in Educational Evaluation*, 18, 365-392.
- Dörner, D. (1976). *Problemlösen und Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Draney, K. & Peres, D. (1998). *Multidimensional modeling of complex science assessment data*. Zugriff am 6. 10. 2012 <http://gse.berkeley.edu/research/BEAR/Publications/multidim.pdf>
- Duit, R. (1995). Zu Rolle der konstruktivistischen Sichtweise in der naturwissenschaftsdidaktischen Lehr- und Lernforschung. *Zeitschrift für Pädagogik*, 41, 905-920.
- Duit, R., Häußler, P. & Prenzel, M. (2001). Schulleistungen im Bereich der naturwissenschaftlichen Bildung. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schule* (S. 169-185). Weinheim und Basel: Hogrefe.
- Dunbar, K. (1993). Concept in a scientific domain. *Cognitive Science*, 17, 397-434.
- Edelmann, W. (2000). *Lernpsychologie*. Kempten: Kösel-Verlag.
- Edwards, J. R. & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5, 155-174.
- Eggert, S. & Bögeholz, S. (2006). Göttinger Modell der Bewertungskompetenz - Teilkompetenz "Bewerten, Entscheiden und Reflektieren" für Gestaltungsaufgaben Nachhaltiger Entwicklung. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 177-197.
- Eid, M. (1999). Lineare Strukturgleichungsmodelle. In B. Strauß, H. Haag & M. Kolb (Hrsg.), *Datenanalyse in der Sportwissenschaft* (S. 427-454). Hofmann: Schorndorf.
- Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah: Lawrence Erlbaum Associates.
- Eschenhagen, D., Kattmann, U. & Rodi, D. (1998). *Fachdidaktik Biologie*. Köln: Aulis Verlag Deubner.

- Felden, H. von (2003). Literacy oder Bildung. Der Literacy-Ansatz der PISA-Studie in bildungstheoretischer Perspektive. In B. Moschner, H. Kiper & U. Kattmann (Hrsg.), *Perspektiven für Lehren und Lernen. PISA 2000 als Herausforderung* (S. 225-240). Battmannsweiler: Schneider Verlag Hohengehren.
- Fischer, H. E., Klemm, K., Leutner, D., Sumfleth, E., Tiemann, R. & Wirth, J. (2003). Naturwissenschaftsdidaktische Lehr-Lernforschung: Defizite und Desiderate. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 179-209.
- Fischer, H. E., Klemm, K., Leutner, D., Sumfleth, E., Tiemann, R. & Wirth, J. (2005). Framework for empirical research on science teaching and learning. *Journal of Science Teacher Education*, 16, 309-349.
- Flora, D. B. & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9, 466-491.
- Ford, D. J., Brickhouse, N. W., Lottero-Perdue, P. & Kittleson, J. (2006). Elementary girls' science reading at home and school. *Science Education*, 90, 270-288.
- Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.
- Frey, A., Hartig, J. & Rupp, A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: theory and practice. *Educational Measurement: Issues and Practice*, 28 (3), 39-53.
- Gentner, D. & Stevens, A. L. (1983). *Mental models*. Hillsdale: Lawrence Erlbaum Association.
- Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41, 867-888.
- Gniewosz, B. (2010). Die Konstruktion des akademischen Selbstkonzeptes. Eltern und Zensuren als Informationsquelle. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42, 133-142.
- Gonzalez, E. J. & Rutkowski, L. (2010). *Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments* (S. 125-156). IEA - ETS Research Institute Monograph 3.

- Gott, R., Duggan, S. & Johnson, P. (1999). What do practising applied scientists do and what are the implications for science education? *Research in Science & Technological Education*, 17, 97-107.
- Gottfredson, L. S. (1997). Intelligence and social policy. *Intelligence*, 24, 13-23.
- Gräber, W. & Bolte, C. (1997). *Toward an understanding of scientific literacy*. Kiel: IPN.
- Gräber, W. & Nentwig, P. (2002). Scientific Literacy - Naturwissenschaftliche Grundbildung in der Diskussion. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung* (S. 7-20). Opladen: Leske + Budrich.
- Gräber, W., Nentwig, P. & Nicolson, P. (2002). Scientific Literacy - von der Theorie zur Praxis. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*. Opladen: Leske + Budrich.
- Gross, L. (2006). Scientific illiteracy and the partisan takeover of biology. *PLoS Biology*, 4, 0680-0683.
- Grube, C. & Mayer, J. (2009, September). *Entwicklung wissenschaftsmethodischer Kompetenzen von Schülerinnen und Schülern in der Sekundarstufe 1 - eine Längsschnittstudie*. Vortrag gehalten auf der Internationalen Tagung der FDdB. Fachsektion Didaktik der Biologie im VBIO, Kiel, Deutschland.
- Grube, C., Möller, A. & Mayer, J. (2007). Dimensionen eines Kompetenzstrukturmodells zum Experimentieren. In H. Bayrhuber, U. Harms, D. Krüger, A. Sandmann, U. Unterbruner, A. Upmeyer zu Belzen & H. Vogt (Hrsg.), *Ausbildung und Professionalisierung von Lehrkräften* (S. 31-34). Kassel: Universität Kassel, FB 18 - Naturwissenschaften.
- Gustafsson, J. & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407-434.
- Hadley, S. T. (1954). A school mark - fact or fancy. *Educational administration and supervision*, 40, 305-312.
- Hammann, M. (2006). Kompetenzförderung und Aufgabenentwicklung. *Der mathematische und naturwissenschaftliche Unterricht*, 59, 85-95.

- Hammann, M. (2007). Das Scientific Discovery as Dual Search-Modell. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden* (S. 187-196). Berlin: Springer.
- Hammann, M., Phan, T. H. & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen. *Zeitschrift für Erziehungswissenschaft, Sonderheft 8*, 33-49.
- Hammann, M. & Prenzel, M. (2008). Ergebnisse des internationalen PISA Naturwissenschaftstests 2006. *Der mathematische und naturwissenschaftliche Unterricht*, 61, 67-74.
- Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. In H. Behrent, H. Dahnke, R. Duit, W. Gräber, M. Komorek, A. Kross & P. Reiska (Hrsg.), *Research in science education - past, present, and future*. Dordrecht: Kluwer Academic Publisher.
- Harlow, A. & Jones, A. (2004). Why students answer TIMSS science test items the way they do. *Research in Science Education*, 34, 221-238.
- Hartig, J., Frey, A. & Jude, N. (2007). Validität. In H. Moosbrugger & A. Keleva (Hrsg.), *Test- und Fragebogenkonstruktion* (S. 135-163). Berlin: Springer.
- Helgeson, S. L. (1994). Research on problem solving: middle school. In D. L. Gabel (Hrsg.), *Handbook of research in science teaching and learning* (S. 248-268). New York: Macmillan.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+R, Manual* (3. Auflage). Göttingen: Beltz-Test.
- Helmke, A. (2000). TIMSS und die Folgen: Der weite Weg von der externen Evaluation der Schulleistungen zur Verbesserung des Lehrens und Lernens. In U. P. Trier (Hrsg.), *Bildungswirksamkeit zwischen Forschung und Politik* (S. 135-164). Bern und Chur: Rüegger.
- Hodson, D. (1992). In search of a meaningful relationship: an exploration of some issues relating to integration in science and science education. *International Journal of Science Education*, 14, 541-562.

- Hodson, D. (2003). Time for action: science education for an alternative future. *International Journal of Science Education*, 25, 645-670.
- Holbrook, J. & Rannikmae, M. (2009). The meaning of scientific literacy. *International Journal of Environmental & Science Education*, 4, 275-288.
- Höttecke, D. (2001). Die Vorstellungen von Schülern und Schülerinnen von der "Natur der Naturwissenschaften". *Zeitschrift für Didaktik der Naturwissenschaften*, 7, 7-23.
- Hülür, G., Wilhelm, O. & Robitzsch, A. (2011). Multivariate Veränderungsmodelle für Schulnoten und Schülerleistungen in Deutsch und Mathematik. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43(4), 173-185.
- Hurd, P. D. (1998). Scientific literacy: new minds for a changing world. *Science Education*, 82, 407-416.
- Ingenkamp, K. (1981). *Wert und Wirkung von Beurteilungsverfahren*. Weinheim und Basel: Beltz.
- Ingenkamp, K. (Hrsg.). (1995). *Die Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz.
- Jansen, M., Schroeders, U. & Lüdtke, O. (in Revision). Academic self-concept in science: multidimensionality, relations to achievement measures, and sex differences.
- Jong, T. De & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31, 105-113.
- Kane, M. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice*, 21 (1), 31-41.
- Kano, Y. & Harada, A. (2000). Stepwise variable selection in factor analysis. *Psychometrika*, 65, 7-22.
- Kattmann, U. (2003a, September). Scientific Literacy und kumulatives Lernen im Biologieunterricht und darüber hinaus - Ein Beitrag zur Bildungsdiskussion nach PISA 2000. In A. Bauer & H. Bayrhuber (Hrsg.), *Entwicklung von Wissen und Kompetenzen. Internationale Tagung der Sektion Biologiedidaktik im VDBiol* (S. 99-103). Kiel: IPN.

- Kattmann, U. (2003b). Vom Blatt zum Planeten. Scientific Literacy und kumulatives Lernen im Biologieunterricht und darüber hinaus. In B. Moschner, H. Kiper & U. Kattmann (Hrsg.), *PISA 2000 als Herausforderung* (S. 115-137). Baltmannsweiler: Schneider Hohengehren.
- Kauertz, A. (2008). *Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben*. Berlin: Logos.
- Keiser, K. K., Nelson, J. E., Norris, N. A. & Szyszkiewicz, S. (1998). *NAEP 1996 science cross-state data compendium for the grade 8 assessment*. Washington DC: NCES.
- Klahr, D. (2000). *Exploring science. The cognition and development of discovery processes*. Massachusetts: Massachusetts Institute of Technology.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., . . . Stanat, P. (2010). *PISA 2009. Bilanz nach einem Jahrzehnt*. Münster: Waxmann.
- Klieme, E., Funke, J., Leutner, D., Reimann, P. & Wirth, J. (2001). Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate einer Schulleistungsstudie. *Zeitschrift für Pädagogik*, 47, 179-200.
- Klieme, E., Hartig, J. & Rauch, D. (2008). The concept of competence in education. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 3-22). Ashland: Hogrefe & Huber Publishers.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 52, 876-903.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Klos, S., Henke, C., Kieren, C., Walpuski, M. & Sumfleth, E. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen - zwei verschiedene Kompetenzen. *Zeitschrift für Pädagogik*, 54, 304-321.

- Kultusministerkonferenz. (2004). Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie. Beschluss der Kultusministerkonferenz vom 01.12.1989 i. d. F. vom 05.02.2004. Zugriff am 6. 10. 2012 http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Biologie.pdf
- Kultusministerkonferenz. (Hrsg.). (2005a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. München, Neuwied: Wolters Kluwer Deutschland.
- Kultusministerkonferenz. (Hrsg.). (2005b). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. München, Neuwied: Wolters Kluwer Deutschland.
- Kultusministerkonferenz. (Hrsg.). (2005c). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. München, Neuwied: Wolters Kluwer Deutschland.
- Kultusministerkonferenz. (Hrsg.). (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Wolters Kluwer Deutschland.
- Kultusministerkonferenz. (2011). Schüler, Klassen, Lehrer und Absolventen 2001 bis 2010. In KMK (Hrsg.), *Statistische Veröffentlichungen der Kultusministerkonferenz. Dokumentation Nr. 195*. Zugriff am 6. 10. 2012 http://www.kmk.org/fileadmin/pdf/Statistik/Dok_195.pdf
- Köller, O. (2008). Bildungsstandards - Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. *Zeitschrift für Pädagogik*, 54, 163-172.
- Köller, O., Baumert, J. & Bos, W. (2001). TIMSS Third International Mathematics and Science Study: Dritte internationale Mathematik- und Naturwissenschaftsstudie. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 270-284). Weinheim und Basel: Hogrefe.
- Koppelt, J. & Tiemann, R. (2008). Modellierung dynamischer Problemlösekompetenz im Chemieunterricht. In D. Höttecke (Hrsg.), *Kompetenzen, Kompetenzmodelle. Kompetenzentwicklung. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Essen 2007* (S. 362-364). Münster: LIT-Verlag.
- Krathwohl, D. R. & Bloom, B. S. (1975). *Taxonomie von Lernzielen im affektiven Bereich*. Weinheim und Basel: Beltz.

- Krüger, D. (2007). Die Conceptual Change-Theorie. In D. Krüger & H. Vogt (Hrsg.), *Handbuch der Theorien in der biologiedidaktischen Forschung* (S. 81-92). Berlin: Springer.
- Kuglemeyer, C. & Schecker, H. (2007). PISA 2000 bis 2006 - Ein Vergleich anhand eines Strukturmodells für naturwissenschaftliche Aufgaben. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 199-220.
- Lau, S. & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment*, 8, 139-162.
- Laugsch, R. C. (2000). Scientific literacy: a conceptual overview. *Science Education*, 84, 71-94.
- Laukenmann, M., Bleicher, M., Fuß, S., Gläser-Zikuda, M., Mayring, P. & Von Rhöneck, C. (2000). Eine Untersuchung zum Einfluss emotionaler Faktoren auf das Lernen im Physikunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 6, 139-155.
- Lavonen, J., Krzywacki, H. & Koistinen, L. (2012). Construction of items for national level assessment in Finnish compulsory school physics without pre-defined learning outcomes. In S. Bernholt, N. Knut & P. Nentwig (Hrsg.), *Making it tangible. Learning outcomes in science* (S. 477-499). Münster: Waxmann.
- Layton, D. (1973). *Science for the people*. London: Allen and Unwin.
- Lee, O. (1997). Scientific literacy for all: what is it, and how can we achieve it? *Journal of Research in Science Teaching*, 34, 219-222.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of nature of science questionnaire: toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39, 497-521.
- Lenneberg, E. H. (1972). *Biologische Grundlagen der Sprache*. Frankfurt am Main: Suhrkamp.
- Lenzen, D. (Hrsg.). (2001). *Pädagogische Grundbegriffe. Band 2. Jugend bis Zeugnis*. Reinbek bei Hamburg: Rowohlt.
- Leucht, M. (2010). *Diagnostik rezeptiver Kompetenzen in der ersten Fremdsprache Englisch auf Basis des Gemeinsamen Europäischen Referenzrahmens für Sprachen* [Monographie]. Dissertation, Freie Universität Berlin.

- Leutner, D. (1997). Instruktions-Design im Spannungsfeld von Behaviorismus, Kognitionspsychologie und Konstruktivismus. In H. Bayrhuber, U. Gebhard, K.-H. Gehlhaar, D. Graf, H. Gropengießer, U. Harms, U. Kattmann, R. Klee & J. C. Schletter (Hrsg.), *Biologieunterricht und Lebenswirklichkeit* (S. 79-97). Kiel: IPN.
- Lewis, S. E., Shaw, J. L. & Heitz, J. O. (2009). Attitude counts: self-concept and success in general chemistry. *Journal of Chemical Education*, 86, 744-749.
- Lintorf, K. (2012). *Wie vorhersagbar sind Grundschulnoten? Prädiktionskraft individueller und kontextspezifischer Merkmale*. Wiesbaden: VS Verlag.
- Lissitz, R. W. (2009). *The concept of validity. Revisions, new directions, and applications*. Charlotte: Information Age Publishing.
- Lissitz, R. W. & Samuelson, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educational Researcher*, 36, 437-448.
- Lüdtke, O., Köller, O., Artelt, C., Stanat, P. & Baumert, J. (2002). Eine Überprüfung von Modellen zur Genese akademischer Selbstkonzepte: Ergebnisse aus der PISA-Studie. *Zeitschrift für Pädagogische Psychologie*, 16, 151-164.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werte in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58, 103-117.
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten. Ansätze zur Bestimmung der Reliabilität und Übereinstimmung von Schülerwahrnehmungen. *Zeitschrift für Pädagogische Psychologie*, 20, 85-96.
- Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage Publications.
- Lukesch, H. (2001). *Psychologie des Lernens und Lehrens*. Regensburg: Roderer.
- Maltby, J., Day, L. & Macaskill, A. (2011). *Differentielle Psychologie, Persönlichkeit und Intelligenz*. München: Pearson Education International.
- Markowitsch, H. J. (1992). *Neuropsychologie des Gedächtnisses*. Göttingen: Hogrefe.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280-295.

- Marsh, H. W. (1990). The structure of academic self-concept: the Marsh/Shavelson model. *Journal of Educational Psychology, 82*, 623-636.
- Marsh, H. W. (1992). Content specificity of relations between academic achievement and academic self-concept. *Journal of Educational Psychology, 84*, 35-42.
- Marsh, H. W. (2002). *Causal ordering of academic self-concept*. Paper presented at the Self-concept Enhancement and Learning Facilitation (SELF) Research Centre International Conference, Sydney. Zugriff am 6. 10. 2012 http://www.self.ox.ac.uk/Conferences/2002_CD_Marsh1.pdf
- Marsh, H. W. & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*, 133-163.
- Marsh, H. W. & Hau, K.-T. (2003). Big-fish-little-pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist, 58*, 364-376.
- Marsh, H. W. & Köller, O. (2004). Unification of theoretical models of academic self-concept/achievement relations: reunification of east and west German school systems after the fall of the Berlin Wall. *Contemporary Educational Psychology, 29*, 264-282.
- Marsh, H. W., Lüdtke, O., Trautwein, U. & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling, 16*, 191-225.
- Marsh, H. W. & Shavelson, R. J. (1985). Self-concept: Its multifaceted-hierarchical structure. *Educational Psychologist, 20*, 107-123.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: construct validation and relations to well-being and achievement. *Journal of Personality, 74*, 403-456.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Development, 76*, 397-416.

- Martin, M. O., Mullis, I. V. S. & Foy, P. (2008). *TIMSS 2007 International science report*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Martin, R. (2011, September). *New aspects related to the theoretical framework of problem solving in PISA 2012*. Paper presented at the 14th Biennale Conference of the European Association for Research in Learning and Instruction, Exeter, England.
- Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden* (S. 177-186). Berlin Heidelberg: Springer.
- Mayer, J., Grube, C. & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In U. Harms & A. Sandmann (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik. Band 3: Ausbildung und Professionalisierung von Lehrkräften. Internationale Tagung der Fachsektion Didaktik der Biologie im VBIO* (S. 63-79). Innsbruck: StudienVerlag.
- McDonald, R. P. (1999). *Test theory. A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, Senatsverwaltung für Bildung, Jugend und Sport & Ministerium für Bildung Wirtschaft und Kultur des Landes Mecklenburg-Vorpommern. (2004). *Rahmenlehrplan Grundschule. Sachunterricht*. Berlin: Wissenschaft und Technik Verlag.
- Minkee, K., Lavonen, J. & Ogawa, M. (2009). Experts' opinions on the high achievement of scientific literacy in PISA 2003: a comparative study in Finland and Korea. *Eurasia Journal of Mathematics, Science & Technology Education*, 5, 379-393.
- Mintzes, J., Wandersee, J. & Novak, J. (1997). *Teaching science for understanding*. San Diego: Academic Press.

- Möller, A., Grube, C., Hartmann, S. & Mayer, J. (2009, März). *Increase of inquiry competence: a longitudinal large-scale assessment of students' performance from grade 5 to 10*. Paper presented at the Annual International Conference of the National Association for Research in Science Teaching, Philadelphia, USA.
- Möller, A., Grube, C. & Mayer, J. (2007). Kompetenzniveaus der Erkenntnisgewinnung bei Schülerinnen und Schülern der Sekundarstufe 1. In H. Bayrhuber, U. Harms, D. Krüger, A. Sandmann, U. Unterbruner, A. Upmeyer zu Belzen & H. Vogt (Hrsg.), *Ausbildung und Professionalisierung von Lehrkräften. Internationale Tagung der Fachgruppe Biologiedidaktik im VBIO*. Kassel: Universität Kassel.
- Möller, A., Hartmann, S. & Mayer, J. (2010, März). *Differentiation and development of five levels on scientific inquiry skills: a longitudinal assessment of biology students in grade 5 to 10*. Paper presented at the International Conference of the National Association of Research in Science Teaching (NARST), Philadelphia, USA.
- Möller, A., Hartmann, S. & Mayer, J. (2009, September). *Modellierung von Niveaus naturwissenschaftsmethodischer Kompetenzen bei Schülerinnen und Schülern der Jahrgangsstufe 5-10*. Vortrag gehalten auf der Internationalen Tagung der Fachsektion Didaktik der Biologie (FDdB) im VBIO, Kiel, Deutschland, S. 40-41.
- Möller, J., Streblow, Pohlmann & Köller, O. (2006). An extension of the internal/external frame of references model to two verbal and numerical domains. *European Journal of Psychology of Education*, 21, 467-487.
- Mullis, I. V. S., Martin, M. O., Robitaille, D. F. & Foy, P. (2008). *TIMSS advanced 2008 international report*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A. & Erberer, E. (2005). *TIMSS 2007 assessment frameworks*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Smith, T. A., Garden, R. A., Gregory, K. D., Gonzalez, E. J., . . . O'Connor, K. M. (2003). *Assessment frameworks and specifications 2003*. Boston: International Study Center, Lynch School of Education, Boston College.

- Muthén, L. K. & Muthén, B. O. (1998-2010). *Mplus user's guide* (Sixth ed.). Los Angeles, CA: Muthén & Muthén.
- National Center for Education Statistics. (2011). *The nation's report card: science 2009*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- National Center for Education Statistics. (2012). *The nation's report card: science 2011*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- National Research Council. (2006). *Systems for state science assessment*. Washington, DC: The National Academies Press.
- Nehring, A., Nowak, K., Tiemann, R. & Upmeyer zu Belzen, A. (2011). "VerE-Studie" - Vernetzung der Erkenntnisgewinnung zwischen Chemie- und Biologieunterricht. In D. Höttecke (Hrsg.), *Gesellschaft für Didaktik der Chemie und Physik: Naturwissenschaftliche Bildung als Beitrag zur Gestaltung partizipativer Demokratie* (S. 510-512). Berlin: Lit.
- Neidorf, T. S., Binkley, M. & Stephens, M. (2006). *Comparing science content in the national assessment of educational progress (NAEP) 2000 and Trends in International Mathematics and Science Study (TIMSS) 2003 assessments*. Washington, DC: National Center for Education Statistics.
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, N. J.: Prentice-Hall.
- Niedersächsisches Kultusministerium. (1995). Grundsätze für eine reformpädagogische Gestaltung des naturwissenschaftlichen Unterrichts. *Schulverwaltungsblatt* (S. 294-298).
- Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching*, 44, 908-937.
- Norris, S. P. & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87, 224-240.

- Oelkers, J. (2002). "Wissenschaftliche Bildung": Einige Verunsicherungen in beiden Richtungen. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung* (S. 105-120). Opladen: Leske + Budrich.
- Opwis, K. (1988). Produktionssysteme. In H. Mandl & H. Spada (Hrsg.), *Wissenspsychologie* (S. 74-98). München, Weinheim: Psychologie Verlags Union.
- Organisation for Economic Co-operation and Development. (2000). *Measuring student knowledge and skills. The PISA 2000 assessment of reading, mathematical und science literacy*. Zugriff am 6. 10. 2012 <http://www.pisa.oecd.org/dataoecd/44/63/33692793.pdf>
- Organisation for Economic Co-operation and Development. (2003). *The PISA 2003 assessment framework: mathematics, reading, science and problemsolving knowledge and skills*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2004). *Learning for tomorrow's world. First results from PISA 2003*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2005). *PISA 2003 Technical report*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2007). *PISA 2006 science competencies for tomorrow's world. Volume 1: analysis*. Paris: OECD Publishing.
- Organisation for Economic Co-operation and Development. (2009). *PISA 2006 technical report*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2010). *PISA 2009 results: what students know and can do. Student performance in reading, mathematics and science. Volume I*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2012). *PISA 2009 technical report*. Paris: OECD Publishing.
- Osborne, J. (2007). Science education for the twenty first century. *Eurasia Journal of Mathematics, Science & Technology Education*, 3, 173-184.

- O'Sullivan, C. Y., Weiss, A. R. & Askew, J. M. (1998). *Students learning science: a report on policies and practices in U.S. schools*. Washington, DC: National Center for Educational Statistics.
- Pajares, F., Britner, S. L. & Valiante, G. (2000). Relation between achievement goals and self-beliefs of middle school students in writing and science. *Contemporary Educational Psychology*, 25, 406-422.
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008). Offenheit für Innovation: Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. *Zeitschrift für Pädagogik*, 56, 827-845.
- Perry, M. (1991). Learning and transfer: instructional conditions and conceptual change. *Cognitive Development*, 6, 449-468.
- PISA-Konsortium Deutschland (Hrsg.). (2004). *PISA 2003 - Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- PISA-Konsortium Deutschland (Hrsg.). (2005). *PISA 2003. Der zweite Vergleich der Länder in Deutschland - Was wissen und können Jugendliche?* Münster: Waxmann.
- PISA-Konsortium Deutschland (Hrsg.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.
- PISA-Konsortium Deutschland (2008). *PISA 2006 in Deutschland - Die Kompetenzen der Jugendlichen im dritten Ländervergleich*. Münster: Waxmann.
- Polanyi, M. (1966). *The tacit knowledge*. Garden City: Doubleday.
- Polya, G. (1957). *How to solve it*. Princeton, NJ: University Press.
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accomodation of a scientific conception: toward a theory of conceptual change. *Science Education*, 66, 211-227.

- Prenzel, M., Geiser, H., Langeheine, R. & Lobemeier, K. (2003). Das naturwissenschaftliche Verständnis am Ende der Grundschule. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 143-182). Waxmann: Münster.
- Prenzel, M., Häußler, P., Rost, J. & Senkbeil, M. (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen. *Unterrichtswissenschaft*, 30, 120-135.
- Prenzel, M., Walter, O. & Frey, A. (2007). PISA misst Kompetenzen. Eine Replik auf Rindermann (2006): Was messen internationale Schulleistungsstudien? *Psychologische Rundschau*, 59, 128-136.
- Prinz, W. (1983). *Wahrnehmung und Tätigkeitssteuerung*. Springer: Heidelberg.
- Reif, F. (1987). Instructional Design, cognition, and technology: applications to the teaching of science concepts. *Journal of Research in Science Teaching*, 24, 309-324.
- Riemeier, T. (2007). Moderater Konstruktivismus. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenden und Doktoranden* (S. 69-79). Berlin: Springer.
- Rieß, W. & Mischo, C. (2007). *Evaluationsbericht "Bildung für nachhaltige Entwicklung (BNE) an weiterführenden Schulen in Baden-Württemberg"*. Maßnahme Lfd. 15 im Aktionsplan Baden-Württemberg. Zugriff am 6. 10. 2012 http://www2.um.baden-wuerttemberg.de/servlet/is/43211/Evaluationsbericht_BNE_2007_Kurzfassung.pdf?command=downloadContent&filename=Evaluationsbericht_BNE_2007_Kurzfassung.pdf
- Rindermann, H. (2006). Was messen Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Grundfähigkeiten, Wissen oder allgemeine Intelligenz? *Psychologische Rundschau*, 57(2), 69-86.
- Rittle-Johnson, B., Siegler, R. S. & Wagner Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: an interative process. *Journal of Educational Psychology*, 93, 346-362.

- Rittle-Johnson, B. & Wagner Alibali, M. (1999). Conceptual and procedural knowledge of mathematics: does one lead to the other? *Journal of Educational Psychology*, 91, 175-189.
- Roberts, R. (2001). Procedural understanding in biology: the 'thinking behind the doing'. *Journal of Biological Education*, 53, 113-117.
- Roberts, R. & Gott, R. (1999). Procedural understanding: its place in the biology curriculum. *School Science Review*, 81(294), 19-25.
- Roberts, R. & Gott, R. (2000). Procedural understanding in biology: how is it characterised in texts? *School Science Review*, 82(298), 83-91.
- Roberts, R. & Gott, R. (2004). A written test for procedural understanding: a way forward for assessment in the UK science curriculum. *Research in Science & Technological Education*, 22, 5-21.
- Rolfhus, E. L. & Ackerman, P. L. (1999). Assessing individual differences in knowledge: knowledge, intelligence, and related traits. *Journal of Educational Psychology*, 91, 511-526.
- Rost, D. H. (2009). *Intelligenz. Fakten und Mythen*. Weinheim, Basel: BeltzPVU.
- Rost, J., Prenzel, M., Carstensen, C. H., Senkbeil, M. & Groß, K. (Hrsg.). (2004). *Naturwissenschaftliche Bildung in Deutschland. Methoden und Ergebnisse von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ruffman, T., Perner, J., Olson, D. R. & Doherty, M. (1993). Reflecting on scientific thinking: children's understanding of the hypothesis-evidence relation. *Child Development*, 64, 1617-1636.
- Rumelhart, D. E. & McClelland, J. L. (1986). *Parallel distributed processing. Explorations on the microstructure of cognition* (Vol. 2). Cambridge: MIT Press.
- Rychen, D. S. (2001). Introduction. In D. S. Rychen & L. Hersh Salganik (Hrsg.), *Defining and selecting key competencies* (S. 1-15). Göttingen: Hogrefe.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Salewski, C. & Renner, B. (2009). *Differentielle und Persönlichkeitspsychologie*. München: Ernst Reinhardt.

- Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Schneider, M., Rittle-Johnson, B. & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, 46, 1525-1538.
- Schneider, M. & Stern, E. (2009). The inverse relation of addition and subtraction: a knowledge integration perspective. *Mathematical Thinking and Learning*, 11, 92-101.
- Schneider, M. & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: a multimethod approach. *Developmental Psychology*, 46, 178-192.
- Schneider, W., Schlagmüller, M. & Ennemoser, M. (2007). *Lesegeschwindigkeits- und -verständnistest für die Klassenstufen 6 -12*. Göttingen: Hogrefe.
- Schnotz, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Koheränzbildung bei Wissenserwerb mit Texten*. Weinheim: Beltz PVU.
- Schön, L.-H., Kramer, J. & Butter, N. (2010). Impulse für die Lehrerbildung. *humboldt-spektrum*, 1-2, 1-5.
- Schroeders, U., Buchholtz, N., Formazin, M. & Wilhelm, O. (in Druck). Modality specificity of comprehension abilities in sciences. *European Journal of Psychological Assessment*.
- Schütte, K., Frenzel, A., Asseburg, R. & Pekrun, R. (2007). Schülermerkmale, naturwissenschaftliche Kompetenz und Berufserwartung. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten Vergleichsstudie* (S. 125-146). Münster: Waxmann.
- Seel, N. M. (2003). *Psychologie des Lernens*. München: Ernst Reinhardt GmbH & Co KG.
- Seidel, T., Prenzel, M., Wittwer, J. & Schwindt, K. (2007). Unterricht in den Naturwissenschaften. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 147-180). Münster: Waxmann.
- Senatsverwaltung für Bildung, Jugend und Sport. (2006). *Rahmenlehrplan für die Sekundarstufe 1*. Berlin: Oktoberdruck Berlin.

- Senkbeil, M., Rost, J., Carstensen, C. H. & Walter, O. (2005). Der nationale Naturwissenschaftstest PISA 2003. Entwicklung und empirische Überprüfung eines zweidimensionalen Facettendesigns. *Empirische Pädagogik*, 19, 166-189.
- Shamos, M. H. (1995). *The myth of scientific literacy*. New Brunswick: Rutgers University Press.
- Shamos, M. H. (2002). Durch Prozesse ein Bewußtsein für die Naturwissenschaften entwickeln. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung* (S. 45-68). Opladen: Leske + Budrich.
- Shavelson, R. J., Baxter, G. P. & Pine, J. (1991). Performance assessment in science. *Applied Measurement*, 4, 347-362.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Review of Educational Research*, 46, 407-411.
- Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M. & Riuz-Primo, M. A. (2008). On the impact of curriculum-embedded formative assessment on learning: a collaboration between curriculum and assessment developers. *Applied Measurement in Education*, 21, 390-402.
- Shen, B. S. P. (1975). Science literacy. *American Scientist*, 63, 265-268.
- Shortland, M. (1988). Advocating science: literacy and public understanding. *Impact of Science on Society*, 38, 305-316.
- Solano-Flores, G. & Shavelson, R. J. (1997). Development of performance assessments in science: conceptual, practical, and logistical issues. *Educational Measurement: Issues and Practices*, 16 (3), 13-22.
- Spearman, C. (1904). "General Intelligence", objectively determined and measured. *American Journal of Psychology*, 15, 201-292.
- Spearman, C. (1927). *The abilities of man: their nature and measurement*. New York: Macmillan.

- Staatsinstitut für Schulqualität und Bildungsforschung München. (2003). *Lehrplan Biologie Realschule 10. Jahrgangsstufe - Bayern*. Zugriff am 6. 10. 2012 <http://www.isb.bayern.de/isb/download.aspx?DownloadFileID=b4e7bb42889dba9a7659f0c38ed9262b>
- Staatsinstitut für Schulqualität und Bildungsforschung München. (2004). *Lehrplan Biologie Gymnasium 10. Jahrgangsstufe - Bayern*. Zugriff am 6. 10. 2012 <http://www.isb-gym8-lehrplan.de/contentserv/3.1.neu/g8.de/index.php?StoryID=26227>
- Steiner, G. (2001). Lernen als Wissenserwerb. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 163-202). Weinheim: PVU.
- Stemmler, G., Hagemann, D., Amelang, F. & Bartussek, D. (2011). *Differentielle Psychologie und Persönlichkeitsforschung*. Stuttgart: W. Kohlhammer GmbH.
- Strobert, G. (2009). Determining validity in national curriculum assessments. *Educational Research*, 51(2), 161-179.
- Süß, H.-M. (2007). Eine Intelligenz - Viele Intelligenzen? Neuere Intelligenztheorien im Widerstreit - 2. Teil. *news&science. Begabtenförderung und Begabtenforschung*, 16, 27-34.
- Taboada, A. & Rutherford, V. (2011). Developing reading comprehension and academic vocabulary for English language learners through science content: a formative experiment. *Reading Psychology*, 32, 113-157.
- Tamir, P. & Doran, R. L. (1992). Conclusions and discussion of findings related to practical skills testing in science. *Studies in Educational Evaluation*, 18, 393-406.
- Tamir, P., Doran, R. L. & Chye, Y. O. (1992). Practical skills testing in science. *Studies in Educational Evaluation*, 18, 263-275.
- Tamir, P., Stavy, R. & Ratner, N. (1998). Teaching science by inquiry: assessment and learning. *Journal of Biological Education*, 33, 27-31.
- Tent, L., Fingerhut, W. & Langfeldt, H.-P. (1976). *Quellen des Lehrerurteils: Untersuchungen zur Aufklärung der Varianz von Schulnoten*. Weinheim und Basel: Beltz.
- Thiel, F. (2007). Stichwort: Umgang mit Wissen. *Zeitschrift für Erziehungswissenschaft*, 10, 153-169.

- Thomas, J. W., Bol, L., Warkentin, R. W., Wilson, M., Strage, A. & Rohwer, W. D. (1993). Interrelationships among students' study activities, self-concept of academic ability, and achievement as a function of characteristics of high-school biology courses. *Applied Cognitive Psychology*, 7, 499-532.
- Thorndike, E. L., Bregman, E. O., Cobb, M. V. & Woodward, E. (1927). *The measurement of intelligence*. New York: Teachers College, Columbia University.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Thurstone, L. L. & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Chicago: University of Chicago Press.
- Tiemann, R., Koppelt, J. & Nehring, A. (2011). Empirische Fundierung chemiedidaktischer Forschung - ein Beitrag zum kompetenztheoretischen Ansatz der Problemlöseforschung. In H. Bayrhuber, U. Harms, B. Muszynski, B. Ralle, M. Rothnagel, L.-H. Schön, H. J. Vollmer & H.-G. Weigand (Hrsg.), *Empirische Fundierung in den Fachdidaktiken*. Münster: Waxmann.
- Trautwein, U. & Baeriswyl, F. (2007). Wenn leistungsstarke Klassenkameraden ein Nachteil sind. *Zeitschrift für Pädagogische Psychologie*, 21, 119-133.
- Treviranus, G. R. (1802). *Biologie, oder Philosophie der lebenden Natur für Naturforscher und Aerzte* (Vol. Band 1). Göttingen: Johann Friedrich Röwer.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Hrsg.), *Organization of memory* (S. 382-404). New York: Academic Press.
- Turmo, A. (2004). Scientific literacy and socio-economic background among 15-year-olds – A nordic perspective. *Scandinavian Journal of Educational Research*, 48, 287-305.
- VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Hrsg.), *Foundation of cognitive science* (S. 527-579). Cambridge, MA: Massachusetts Institute of Technology Press.
- Vernon, P. E. (1950). *The structure of human abilities*. Oxford: Wiley.
- Upmeyer zu Belzen, A. & Krüger, D. (2010). Modellkompetenz im Biologieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 41-57.

- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249-294.
- Watson, F. & Cohen, I. (1952). *General education in science*. Cambridge, MA: Harvard University Press.
- Weber, H. & Ramseyer, T. (2012). *Differentielle Psychologie - Persönlichkeitsforschung*. Göttingen: Hogrefe.
- Weinert, F. E. (1992). *Wie groß ist der Einfluß der Schule auf die geistige Entwicklung der Schüler, - und wie groß könnte er sein?, Vortrag an der Universität des Saarlandes Saarbrücken, 16. Juni 1992*.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schule* (S. 17-31). Weinheim und Basel: Beltz.
- Weinert, F. E. & Helmke, A. (1997). *Entwicklung im Grundschulalter*. Weinheim: Beltz PVU.
- Wellington, J. & Osborne, J. (2001). *Language and literacy in science education*. Buckingham: O University Press.
- Westby, C. & Torres-Velásquez, D. (2000). Developing scientific literacy. *Remedial and Special Education*, 21, 101-110.
- White, R. W. (1959). Motivation reconsidered. The concept of competence. *Psychological Review*, 66, 297-333.
- Whitehead, A. N. (1929). *The aims of education*. New York: Macmillan.
- Wilson, M. & Boeck, P. De (2004). Descriptive and explanatory item response models. In P. De Boeck & M. Wilson (Hrsg.), *Explanatory item response models. A generalized linear and nonlinear approach* (S. 43-74). New York: Springer.
- Winkelmann, H. (2009). *Validierung der länderübergreifenden Bildungsstandards für mathematische Kompetenzen im Primarbereich* [Monographie]. Dissertation, Humboldt Universität zu Berlin.
- Winick, D. M., Avallone, A. P., Smith, C. E. & Crovo, M. (2008). *Science framework for the 2009 National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board.

- Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer Science+Business Media, LLC.
- Wong, S. L. & Hodson, D. (2009). From the horse's mouth: What scientists say about scientific investigation as scientific knowledge. *Science Education*, 93, 109-130.
- Wygotzki, L. S. (1964). *Denken und Sprechen*. Stuttgart: Fischer.
- Zanobini, M. & Usai, C. (2002). Domain-specific self-concept and achievement motivation in the transition from primary to low middle school. *Educational Psychology*, 22, 203-217.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27, 172-223.
- Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. W. Lissitz (Hrsg.), *The concept of validity. Revisions, new directions, and applications*. Charlotte: Information Age Publishing.

Selbstständigkeitserklärung

Hiermit versichere ich, Nele Nicole Kampa, dass ich die vorliegende Dissertation selbstständig verfasst, ohne unzulässige Hilfe Dritter und ohne Nutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Die aus fremden Quellen direkt oder indirekt übernommenen Inhalte sind unter Angaben der Quellen kenntlich gemacht. Dies gilt auch für bildliche Darstellungen sowie für Quellen aus dem Internet.

Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Berlin, Oktober 2012

Nele Nicole Kampa

